Application Of Joint Intensity Algorithms To The Registration Of Emission Tomography And Anatomical Images

by

Yiu Hon LAU

A thesis
submitted for the Degree of
Doctor of Philosophy

University of Technology, Sydney
July 2004

CERTIFICATE OF AUTHORSHIP / ORIGINALITY

I certify that this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature of Candidate

Production Note:
Signature removed prior to publication.

Acknowledgement

I would like to express my grateful appreciation to Prof. Michael Braun and Prof. Brian F. Hutton who have touched my life by offering their support and assistance to make this thesis possible. They were the driving force in keeping me going when I almost gave up the study. They encouraged me in my area of research and inspired me to submit papers for local and international conferences. There is no doubt that they contributed considerably to the readability and content of this thesis, for which I wish to thank them. I recall the memorable workshop we attended together at Utrecht in Netherlands. They motivated me and made me feel that my work was valued. Prof. Braun introduced me to relevant digital image processing algorithms, and aspects of CT and MR physics. His experience in the field, his many ideas and suggestions have been extremely valuable and instructive. Prof. Hutton introduced me to image reconstruction techniques, the MCAT phantom, IDL programming language and gamma camera physics. What I learnt about SPET cameras, their operation, data acquisition and processing has proved to be of great benefit to this study. Prof. Hutton's expertise in nuclear medicine together with his leadership was invaluable to the research group at the Westmead Hospital. I appreciate the supportive and friendly atmosphere in the group, enhancing scientific interaction among us. Discussions with my supervisors helped me to keep both feet on the ground, which was not always easy.

Special thanks are due to my colleague, Mr. David Skerrett. It has been a privilege to know and work with him over the past few years. He is a senior medical physicist with a special talent in computer networks. He taught me much about UNIX. With his assistance, my life in the computer laboratory was made easier. He also gave me support in the use of software applications, UNIX workstations, networks, MO disks and drives, FTP and emails… and he generously gave of his personal time. I enjoyed talking with him especially when we walked together to the car park after a day of hard work. He kindly proofread the entire thesis. I wish him all the best.

Acknowledgment is also due Dr. Leighton Barnden, Chief Medical Physicist at the Department of Nuclear Medicine of Queen Elizabeth Hospital in Adelaide for providing me with advice on my registration software. He supplied clinical data for validation of the automatic retrospective multimodality image registration package *MIR*. I also thank

Dr. Lennart Thurfjell, then of Uppsala University, for the enlightening discussion about image registration during his visit to Sydney. He gave me the inspiration for the implementation of the registration algorithm. Thanks also to Dr. D.K. Chung and Dr. R. Howman-Giles of the Department of Nuclear Medicine at the Children's Hospital at Westmead in Sydney for critically evaluating *MIR* in clinical practice. They have offered me the clinical data and the opportunity to perform this research.

I thank Rochelle Hatton and Seu Som for invaluable discussions and for the enjoyable atmosphere that made the research pleasant and productive. It was a great pleasure to work with them. I also wish to thank Dr. George Larcos, Dr. David Farlow, Dr. Simon Gruenewald, Dr. Kevin Choong, Dr. Lee Collins and Mr. George Bonovas for supporting this research at the Westmead Hospital. I also wish to acknowledge the contributions of Mr. Erin McKay at St. George Hospital, Sydney for providing SPET and CT images, Dr. S. Meikle at the Royal Prince Alfred Hospital, Sydney, for providing the OSEM reconstruction source code, and Dr. George Zubal of Yale University for providing the Zubal phantom. I acknowledge the financial support provided by the Australian government via the Australian Postgraduate Award as well as the Department of Medical Physics, the Westmead Hospital, and the Department of Applied Physics, University of Technology, Sydney, for assistance with conference expenses.

Finally, from the bottom of my heart, I thank my "better half", Carol, who was always there to support me, along with my kids: Eppie and Anisa. Without their support it would not have been possible for me to finish this study.

The work reported in this dissertation was carried out in the Department of Applied Physics at the University of Technology in Sydney, and the Department of Medical Physics & Department of Nuclear Medicine and Ultrasound at Westmead Hospital in Sydney.

Table of contents

List of acronyms and mathematical symbols by chapter

Chapter 1

2D	two dimensional
3D	three dimensional
CT	computerized tomography
MR	magnetic resonance
NMR	nuclear magnetic resonance
SPET	single photon emission tomography
PET	positron emission tomography
PACS	picture archiving and communication system
μ	linear attenuation coefficient
$\mathbf{M}$	macroscopic magnetization
$\mathbf{M}_{xy}$	transverse magnetization
$\mathbf{M}_Z$	axial magnetization
$\mathbf{B}$	magnetic field
T_1	longitudinal relaxation time / spin-lattice relaxation time
T_2	transverse relaxation time / spin-spin relaxation time
PD	proton density
γ	gyromagnetic ratio
ω_o	larmor frequency
RF	radio frequency
G	magnetic field gradient
FDG	^{18}F-fluoro-deoxyglucose

Chapter 2

fMR	functional MR
$\mathbf{Q}$	4×4 matrix representing a rigid-body transformation in homogeneous coordinates
$\mathbf{Q}'$	3×3 rigid-body transformation matrix
$\mathbf{T}$	4×4 translational matrix in homogeneous coordinates
$\mathbf{R}$	4×4 rotational matrix in homogeneous coordinates
$\mathbf{S}$	4×4 scaling matrix in homogeneous coordinates
$\mathbf{I}$	Identity transformation
$\mathbf{R}'$	3×3 orthonormal rotation matrix
$\mathbf{t}$	3-element translation vector (t_x, t_y, t_z)
$\boldsymbol{\varphi}$	3-element rotation vector (φ_x, φ_y, φ_z)
$\mathbf{S}'$	3-element scale vector (s_x, s_y, s_z)
$\bar{y}, \bar{x}$	centroid of $\{y_i\}$ and $\{x_i\}$
SVD	singular-value decomposition
$\mathbf{Q}^T$	transpose of a matrix Q
$\mathbf{D}$	diagonal matrix
$\mathbf{U}$, $\mathbf{V}$	orthonormal matrices
$\|x - y\|$	Euclidean distance between x and y

SAD	sum of the absolute differences
SSD	sum of squared differences
$\overline{\mathrm{A}}$	mean intensity in image A
N	number of voxels in the overlapping image volume
F	number of fiducial markers
PV	partial volume interpolation
Φ	similarity measure
M	number of intensity bins
FRE	fiducial registration error
TRE	target registration error

Chapter 3

A	floating image to be aligned with the reference image
A′	transformed floating image (unless otherwise indicated, this image is interpolated and sampled in the domain of the reference image)
B	reference image
x	voxel position expressed in homogeneous coordinates
$a = \mathrm{A}(x)$	voxel intensity of image A at x
$b = \mathrm{B}(x)$	voxel intensity of image B at x
$\widetilde{\Omega}_A$	continuous domain ("field of view") of image A
Ω_A	discrete image domain of A: $\Omega_A \subset \widetilde{\Omega}_A$
Ω_{AB}	discrete overlap of image domains of A′ and B, $\Omega_{AB} = (\widetilde{\Omega}_{A'} \cap \widetilde{\Omega}_B) \cap \Omega_B$
S_b	a subimage of A′ induced by b in B such that $\mathrm{S}_b = \{x \in \Omega_{AB} : \mathrm{B}(x) = b\}$
T	a mapping of A to A′ that includes interpolation
T_θ	a mapping parameterized by transformation parameters θ
θ	a vector of registration parameters $(\varphi_x,\varphi_y,\varphi_z,t_x,t_y,t_z)$
$\mathrm{T}_\theta\mathrm{A}$	transformed image A′
PDF	probability density function
Q	matrix representing a spatial transformation of voxel coordinates, $x' = \mathbf{Q}x$
$\mathrm{A}(\mathbf{Q}x)$	transformed image A′ such that $\mathrm{A}'(x) = \mathrm{T}_\theta\mathrm{A}(x) = \mathrm{A}(\mathbf{Q}\,x)$
x'	location of voxel x transformed by **Q**: $x' = \mathbf{Q}x$
$\mathrm{P}(a)$	probability density function of intensity in image A
$\mathrm{P}(b)$	probability density function of intensity in image B
(a,b)	intensity pair of a joint intensity histogram of images A and B, where a belongs to the range of A and b belongs to the range of B
$\mathrm{P}(a,b)$	joint probability of A and B
$\mathrm{P}(a\|b)$	conditional probability of A given B
H(A)	entropy of an image A
H(A,B)	joint entropy of images A and B
H(A|B)	conditional joint entropy of image A given the knowledge of B
MI(A,B)	mutual information of images A and B
NMI(A,B)	normalized mutual information of images A and B

μ_A	mean intensity of image A
μ_b	mean intensity of a subimage S_b
$\sigma^2(a)$	variance of intensity of image A
$\sigma^2(a \mid b)$	conditional variance of intensity of a subimage S_b given b in B
CR(A,B)	correlation ratio of A and B
SCR(A,B)	symmetric correlation ratio of A and B

Chapter 4

$[n_1,n_2,n_3\|M]$	downsampling of the transformed floating image $T_\theta A$ and the reference image B by a factor of n_1 in the x-direction, n_2 in the y-direction and n_3 in the z-direction, and M is the number of bins in the intensity histogram
^{99m}Tc-HMPAO	technetium-99m hexamethyl-propylene amine oxime (HMPAO)
DE	displacement error measures the average Euclidean displacement over all eight vertices of a bounding box
FLE	fiducial location error

Chapter 5

PMT	photomultiplier tube
NaI (Tl)	sodium iodide crystal doped with about 0.5% of thallium oxide
PHA	pulse height analyzer
PSF	point-spread function
FWHM	full width at half maximum
d	path length in an attenuating medium
$f(x,y)$	radioactivity distribution in Cartesian coordinates
$\hat{f}(x,y)$	estimated radioactivity distribution
FBP	filtered back projection reconstruction
$g(s,\theta)$	slice profile (projection) is defined as the total count detected in a time interval at pixel s from the origin when the detector is at an angular position θ
p	total number of projections
MLEM	maximum likelihood expectation maximization reconstruction
OSEM	ordered Subsets Expectation Maximization reconstruction
N_θ, N	number of projection angles and number of bins per projection
$\boldsymbol{A}$	transition matrix such that $g = \boldsymbol{A}f$
g_i	number of counts in the i^{th} measurement
α, β	collimator constant and collimator scale of a given gamma camera
d	source-detector distance
σ	Gaussian width $\sigma = \alpha + \beta \times d$ used to estimate the PSF of a gamma detector

Chapter 6

θ	rigid body transformation, $\theta(x) = \mathbf{R}x + \mathbf{t}$
$\mathbf{R}$, $\mathbf{t}$	rotational and translational misalignment
θ_{app}	applied transformation

θ_{rec}	recovered transformation
ε	registration error
ε_R	rotational error, $\varepsilon_{\text{R}} = \frac{1}{8}\sum_{i=1}^{8}\left\|(\mathbf{R}_{\text{app}} - \mathbf{R}_{\text{rec}})\, x_{\text{i}}\right\|$
ε_t	translation error, $\varepsilon_{\text{t}} = \left\|\mathbf{t}_{\text{app}} - \mathbf{t}_{\text{rec}}\right\|$
$\bar{\varepsilon}$	mean registration error
AC	attenuation correction
SD	standard deviation

Chapter 7

SSD	sum of squared differences of voxel values between images
$u(x)$	displacement vector at x
$\tilde{u}$	measured displacement vector
T_u	transformation T parameterized by a displacement vector field u
$\mathrm{T}_u\mathrm{A}(x)$	floating image A transformed with respect to a *reference* image B using transformation T parameterized by a displacement vector $u(x)$ where x is the position vector in the reference image such that $\mathrm{A}'(x) = \mathrm{T}_u\mathrm{A}(x) = \mathrm{A}(x - u(x))$
σ_b^2	conditional variance indexed by b
σ^2	variance of A′
i	image voxel index $i \in \{0, 1,..., \mathrm{N}\text{-}1\}$
j	grid nodes index j
ϕ	Gaussian kernal $\phi\left(x_{ij}\right) = \exp\left(-\frac{1}{2}\left\|x_{ij}\right\|^2 / \delta^2\right)$
δ	Gaussian width
$\left\|x_{ij}\right\|$	Euclidean norm of $\left\|x_i - x_j\right\|$ on R^3
$\left(\mathrm{c}_j\right)$	column vector consisting of n coefficients corresponding to n grid nodes such that $u(x_i) = \sum_{j\in Q}\phi\left(x_{ij}\right)\mathrm{c}\left(x_j\right)$
$\left(\phi_{j'j}\right)$	$n \times n$ matrix of Gaussian weights $\phi_{j'j}$
v	residual displacement vector field obtained by convolution $v\left(x_i\right) = \phi\left(x_i\right) * \tilde{v}\left(x_i\right)$
ε	displacement error given by $\varepsilon_i = \left\|u_i - \tilde{u}_i\right\|$
DE	RMS displacement error
ADE	average RMS displacement error

Chapter 8

ε	displacement error given by $\varepsilon_i = \left\|u_i - \tilde{u}_i\right\|$
DE	RMS displacement error
ADE	average RMS displacement error over all images

List of illustrations

List of tables

Appendix

Abstract

In current practice, it is common in medical diagnosis or treatment monitoring for a patient to require multiple examinations using different imaging techniques. Magnetic resonance (MR) imaging and computed tomography (CT) are good at providing anatomical information. Three-dimensional functional information about tissues and organs is often obtained with radionuclide imaging modalities: positron emission tomography (PET) and single photon emission tomography (SPET). In nuclear medicine, such techniques must contend with poor spatial resolution, poor counting statistics of functional images and the lack of correspondence between the distribution of the radioactive tracer and anatomical boundaries. Information gained from anatomical and functional images is usually of a complementary nature. Since the patient cannot be relied on to assume exactly the same pose at different times and possibly in different scanners, spatial alignment of images is needed. In this thesis, a general framework for image registration is presented, in which the optimum alignment corresponds to a maximum of a similarity measure. Particular attention is drawn to entropy-based measures, and variance-based measures. These similarity measures include mutual information, normalized mutual information and correlation ratio which are the ones being considered in this study. In multimodality image registration between functional and anatomical images, these measures manifest superior performance compared to feature-based measures. A common characteristic of these measures is the use of the joint-intensity histogram, which is needed to estimate the joint probability and the marginal probability of the images. A novel similarity measure is proposed, the symmetric correlation ratio (SCR), which is a simple extension of the correlation ratio measure. Experiments were performed to study questions pertaining to the optimization of the registration process. For example, do these measures produce similar registration accuracy in the non-brain region as in the brain? Does the performance of SPET-CT registration depend on the choice of the reconstruction method (FBP or OSEM)? The joint-intensity based similarity measures were examined and compared using clinical data with real distortions and digital phantoms with synthetic distortions. In automatic SPET-MR rigid-body registration applied to clinical brain data, a global mean accuracy of 3.9 mm was measured using external fiducial markers. SCR performed better than mutual information when sparse sampling was used to speed up the registration process. Using the Zubal phantom of the thoracic-abdominal region, SPET projections for

Methylenediphosponate (MDP) and Gallium-67 (^{67}Ga) studies were simulated for 360° data, accounting for noise, attenuation and depth-dependent resolution. Projection data were reconstructed using conventional filtered back projection (FBP) and accelerated maximum likelihood reconstruction based on the use of ordered subsets (OSEM). The results of SPET-CT rigid-body registration of the thoracic-abdominal region revealed that registration accuracy was insensitive to image noise, irrespective of which reconstruction method was used. The registration accuracy, to some extent, depended on which algorithm (OSEM or FBP) was used for SPET reconstruction. It was found that, for roughly noise-equivalent images, OSEM-reconstructed SPET produced better registration than FBP-reconstructed SPET when attenuation compensation (AC) was included but this was less obvious for SPET without AC. The results suggest that OSEM is the preferable SPET reconstruction algorithm, producing more accurate rigid-body image registration when AC is used to remove artifacts due to non-uniform attenuation in the thoracic region. Registration performance deteriorated with decreasing planar projection count. The presence of the body boundary in the SPET image and matching fields of view were shown not to affect the registration performance substantially but pre-processing steps such as CT intensity windowing did improve registration accuracy. Non-rigid registration based on SCR was also investigated. The proposed algorithm for non-rigid registration is based on overlapping image blocks defined on a 3D grid pattern and a multi-level strategy. The transformation vector field, representing image deformation is found by translating each block so as to maximize the local similarity measure. The resulting sparsely sampled vector field is interpolated using a Gaussian function to ensure a locally smooth transformation. Comparisons were performed to test the effectiveness of SCR, MI and NMI in 3D intra- and inter-modality registration. The accuracy of the technique was evaluated on digital phantoms and on patient data. SCR demonstrated a better non-rigid registration than MI when sparse sampling was used for image block matching. For the high-resolution MR-MR image of brain region, the proposed algorithm was successful, placing 92% of image voxels within $\leq$2 voxels of the true position. Where one of the images had low resolution (e.g. in CT-SPET, MR-SPET registration), the accuracy and robustness deteriorated profoundly. In the current implementation, a 3D registration process takes about 10 minutes to complete on a stand alone Pentium IV PC with 1.7 GHz CPU and 256 Mbytes random access memory on board.

Chapter 1

Introduction

The developments of medical imaging apparatus and reconstruction algorithms have advanced greatly over the past two decades. Ionizing radiation has been used to produce images of the interior of the body and in the treatment of cancer. Gamma radiation produced by radionuclides and x-rays are both used to study specific organs and diseases in the body via 2D planar images. The planar projection images are overlaid with internal body structures that cause deterioration in delineating anatomical structure and tissues inside body. With the development of digital computers it was subsequently possible to reconstruct cross-sectional images of body, resulting in computerized tomography. Another important development utilizing non-ionization radiation has been magnetic resonance imaging which has become a major diagnostic imaging tool. In the following sections, a brief introduction of different image modalities is discussed, followed by a short overview of the rationale for image registration and finally defining problems for image registration.

1.1 Historical overview of medical imaging

In 1895 Roentgen discovered x-rays and in 1896 Becquerel discovered radioactivity and very soon afterwards these were applied in medicine [1,2]. The use of x-rays and radioactivity to investigate the body resulted in the development of the field of diagnostic imaging in radiology and nuclear medicine, respectively. X-rays are a form of electromagnetic radiation with a very short wavelength. They are emitted when high-speed electrons are rapidly decelerated. The basis of x-ray imaging is primarily based on the interaction of x-ray with matter to produce contrast between anatomical structure and background. This interaction leads the intensity of the radiation attenuating as it passes through matter. The attenuation depends on the composition of the material. After Roentgens discovery photographic film and the fluorescent screen have been used to display x-ray images. The film is used as a permanent record of a 2D projective image. In a chest x-ray radiograph, as shown in Figure 1.1, the 3D structure of the body is represented by a 2D image. All the planes in the patient that are parallel to the x-ray film are superimposed on top of each other.

The process of creating a 2D projective image from the emissions of radionuclides within the body is quite different from that for x-rays, where radiation is normally applied from an external source while radioactive substances are taken internally. The first gamma camera that was capable of producing a 2D projection image of the radioactive distribution of an administered radionuclide in the body was invented in 1957 by Anger. Because of the overlay of structures of the body, one intrinsic problem of these planar projection images is the reduction of contrast of anatomical structures and the lost of spatial resolution in the direction perpendicular to the image plane.

A further development in the 1970s was the introduction of tomography (both emission and transmission). X-ray computerized tomography (CT) was developed by Godfrey Hounsfield [3] and Allan Cormack. It is of interest to note that the mathematics of reconstruction of a function from its projections had been studied in the early 1900s by Radon. Another important development in diagnostic imaging was nuclear magnetic resonance imaging, realized in the mid 1970s [2]. Modern medical imaging applications were made possible by the development of computer technology and the availability of high quality 3D image data. With the improvements in imaging instrumentation and algorithms for image reconstruction, modern imaging technologies provide a wide range of non-invasive image acquisition capabilities. These modalities can be used to acquire data of human anatomy and function in 3D. Whereas conventional planar imaging results in a two-dimensional projective view through the body, computed tomography produces images of transaxial slices of the body.

The problem of obscuring body structures has been largely eliminated by computed tomography. Furthermore, images produced are in digital format, which facilitates ease of display and analysis. Software must handle more and more complex problems demanded by the ongoing needs of clinical applications, such as those arising in neurology, radiotherapy, cardiology and computer-aided surgery. The development of 3D medical image processing techniques has not occurred in isolation. In large part, it has been due to the increased availability of fast computers and digital networking, enabling a rapid, accurate manipulation of large quantities of data to produce images that are useful to physicians. Picture Archiving and Communication System (PACS) is an acknowledged requirement for the modern radiology department [4,5]. The ease of accessing images from multiple modalities has led to a greater use of image fusion that

drives registration development [4,5]. For routine clinical practice applications, automatic software is required to align medical images. This thesis is concerned with the automatic retrospective spatial alignment of images from different image modalities.

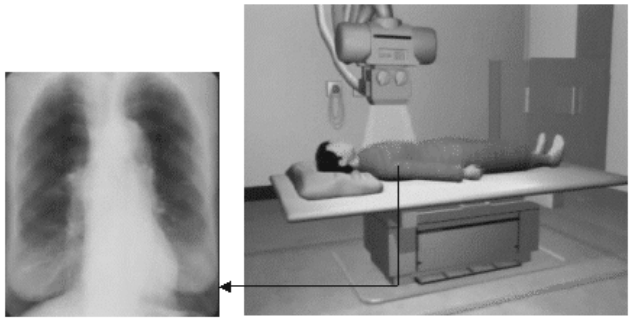


Figure 1.1: An x-ray film. The attenuation information for the patient's chest is superimposed onto the (horizontal) image plane. The output image shows overlying ribs, soft tissue and lungs. (Image courtesy of Curriculum Development Institute of Hong Kong)

1.2 Functional and anatomical medical images

The human body contains many structures overlying other structures within the body. This limits the amount of information that can be obtained from conventional planar images. A significant advance in the use of medical images came with the development of computed tomography. This technique makes it possible to image axially, or view slices at any cross section required. Tomography refers to any process that produces an image of a single plane of the body [6]. When multiple contiguous slices are reconstructed, the stack of these 2D slices makes up a 3D volume. If the projections are formed by radiation transmitted through the body, the technique is referred to as transmission computed tomography (CT), the best known example being x-ray CT. In radionuclide imaging, it is called emission tomography (ET). The method that uses the rotating gamma camera(s) to detect photons is called the single photon ET, or SPET. In positron emission tomography, or PET, rings of detectors detect the annihilation radiation emitted from a patient injected with a positron emitting radionuclide. A further form of tomography, which is known as magnetic resonance or MR, uses entirely different physical phenomena. MR is based on the principles of nuclear magnetic resonance, a spectroscopic technique used by scientists to obtain microscopic chemical and physical information about molecules. A transverse slice for CT, MR and SPET is depicted in Figure 2.1. In the remainder of this section, physical principles underlying each modality are described briefly and their clinical applications are outlined.

1.2.1 Anatomical imaging

Because x-ray CT and MR produce high resolution images with excellent depiction of body tissues, these two image modalities have long been the choice for examination of anatomical information. As the techniques, technology and computer processing power have improved, their applications have multiplied.

X-ray computed tomography

X-ray computed tomography is a medical imaging technique that produces image of transaxial planes through a body. When compared to the conventional 2D radiograph, which is a many planes superimposed on each other, a CT image exhibits significantly improvement in contrast (i.e. the ability to distinguish an object with its background). In x-ray radiography and in CT, the incident x-rays are a form of electromagnetic radiation with a very short wavelength. An x-ray source, such as a rotating anode x-ray tube, is often used to generate x-rays by bombarding a piece of target metal with high-speed electrons. The electrons are produced by thermionic emission from a hot cathode. They are accelerated by a high voltage potential and collide with the target. The x-ray tube used in CT is typically operated in the range of 100 kV to 150 kV. The incident x-ray beam is attenuated by interaction of the x-ray photons with body tissue. Attenuation of x-rays is most easily explained using a narrow (pencil) beam. Assuming a pencil beam of monoenergetic x-rays, the intensity I of the beam after transmission through a small volume of tissue of infinitesimal thickness δx_i and uniform attenuation coefficient μ_i is given by

$$I = I_o e^{-\mu_i \delta x_i}, \qquad (1.1)$$

where I_o is the incident x-ray intensity. The attenuation depends on the composition of the material. Comparing two types of body tissue e.g. muscle and bone, shows the difference well. In general, four processes (coherent scatter, photoelectric effect, Compton scatter and pair production[1]) can result in loss of energy from x-ray beams in tissue. A detailed description of each process is outside the scope of this thesis. Interested readers may refer to [7] for detailed description of these processes.

In a radiograph, the 3D structure of the body is compressed into a 2D image. In this case, overlying structures within the body are superimposed. Because of the superposition,

[1] Does not occur at diagnostic x-ray energy levels.

radiographs do not exhibit high contrast and depth information. Bone and air cavities are easily seen but very little contrast is obtained in soft tissue regions, e.g. a blood vessel surrounded by muscles will not be seen. X-ray computed tomography produces an image of a 2D slice through the body, thus reducing the problem of superposition to the small thickness of the slice. The original CT scanners generated a pencil shaped narrow beam of x-rays from an x-ray tube. An array of several hundred x-ray detectors, are mounted so as to receive the beam as the tube rotates. The x-ray detectors in a CT scanner do not directly produce an image. The detector measures x-ray intensity transmitted through the chosen slice from many directions. Figure 1.2 depicts a simple CT scanner and illustrates acquisition of a large number of projections taken at different angles as the x-ray tube rotates around the patient.

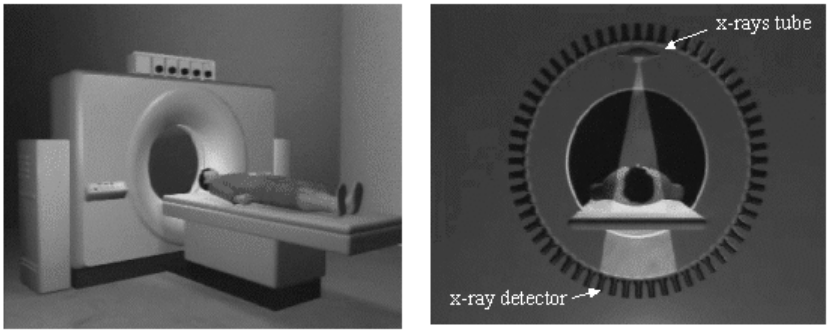


Figure 1.2: An x-ray computed tomography unit. As the x-ray tube rotates, a fan-beam of x-rays is scanned across the patient to acquire a projection that is made up of a large number of rays transmitted through the body. (Image courtesy of Curriculum Development Institute of Hong Kong)

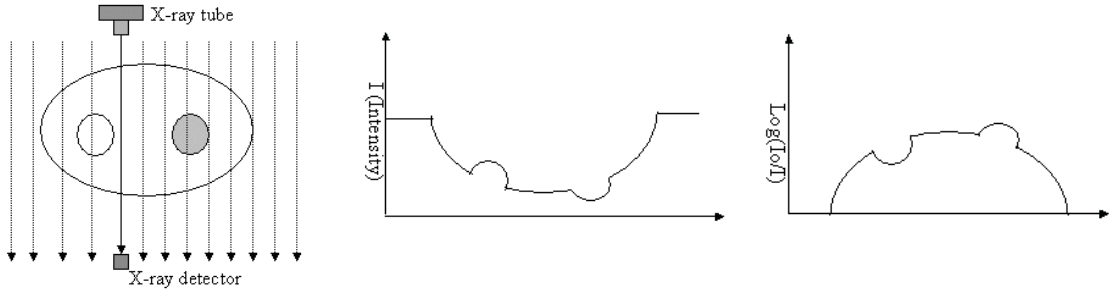


Figure 1.3: A schematics diagram to show a single projection of CT scan. An intensity profile (middle) and an attenuation profile (right) are illustrated.

The emerging (pencil) monoenergetic x-ray beam will have been attenuated by n voxels, each of thickness δx, where the attenuation in each voxel is given by Equation (1.1). The combined attenuation is expressed as

$$\ln\left(\frac{I_o}{I}\right) = \sum_{i=1}^{n} \mu_i \delta x_i \ , \qquad (1.2)$$

where *i* is an index for voxels in the path of the x-rays. If, as is often the case, the path length through each voxel is the same, the logarithm of the relative intensity is proportional to the sum of the attenuation coefficients along the path through the body, hence the term *raysum*. A single parallel projection is shown in Figure 1.3 to illustrate schematically an intensity profile and its corresponding attenuation profile using Equation (1.2). Modern CT scanners use a fan-beam geometry but the parallel beam is shown here for the sake of simplicity. The CT image is reconstructed mathematically (usually using filtered back projection as discussed in Subsection 5.4.1) from a large number of projections of a chosen plane. Each projection is made up of a large number of raysums, acquired electronically using a linear array of solid-state detectors and an x-ray source at different angles around the patient [8]. The in-plane resolution is typically 0.5-2.0 mm and slice thickness is typically 1-10 mm [9]. In a CT image (Figure 1.4), the value of each voxel is related to the local linear attenuation coefficient. The measured attenuation coefficients are normally converted to CT numbers, which express the attenuation coefficients relative to the attenuation coefficient of water. Due to its high contrast and spatial resolution, CT is used routinely in radiology for a wide range of clinical applications that include diagnosis, surgical planning [10,11], prosthesis design [12], stereotactic biopsy planning [13,14] and radiotherapy treatment planning [15,16].

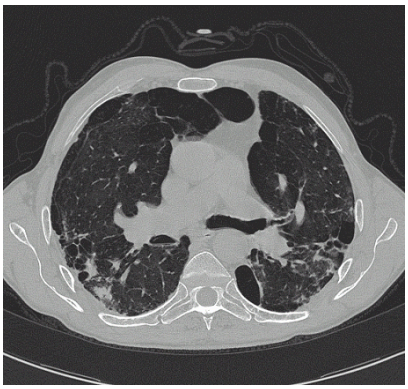


Figure 1.4: A typical CT scan of the thorax region of a body showing high contrast and resolution.

Magnetic resonance imaging

In contrast to x-ray CT using ionizing radiation to produce images of the inside of the human body, MR is free of the hazards associate with ionizing radiation. Over the past twenty years, MR has become one of the most important imaging modalities. Reviews of the relevant magnetic resonance physics and MR imaging principles can be found in [7,17,18,19,20]. According to classical electromagnetism, a spinning sphere of charge

generates a magnetic dipole moment like a microscopic bar magnet. The hydrogen nucleus exhibits a non-zero spin. In a unit volume element corresponding to a voxel of an image, there are many protons, each with an associated magnetic dipole moment, and the net macroscopic magnetization **M** of the volume element is the vector sum of the individual dipole moments. In the absence of an external magnetic field, the spatial orientation of each dipole moment is random and hence **M** = 0. Because of the very high water content in the human body, hydrogen is the most common element in tissues and gives rise to a relatively strong MR signal. For that reason, hydrogen is the mainstay of MR imaging. When a static magnetic field (main field) $\mathbf{B}_0$ is applied, the random microscopic oriented dipoles respond to the force of the external magnetic field that are exerted on them by trying to align with it. However, the alignment is not instantaneous. It takes a finite time to overcome loss of energy from the spin system due to thermal randomization. When $\mathbf{B}_0$ (along the z-axis) is turned on, the net magnetization **M** takes time to grow from zero towards its full longitudinal magnetization value $\mathbf{M}_0$ along the z-axis. The net bulk magnetization **M** are generally not aligned exactly with the external main field $\mathbf{B}_0$, but rather make an angle, and will be forced into a precessional motion. This is analogous to the wobbling of a top as it spins under the influence of the earth's gravitational force. The precession motion of **M** about $\mathbf{B}_0$ can be described by $\frac{d\mathbf{M}(t)}{dt} = \gamma \mathbf{M} \times \mathbf{B}_o$. The precession frequency, known as the Larmor frequency, is proportional to $\mathbf{B}_0$, $\omega_o = \gamma \mathbf{B}_o$. The constant of proportionality is the so-called gyromagnetic ratio γ, which is characteristic of each nuclear species. Longitudinal magnetization along the main magnetic field cannot be measured directly. If a radio frequency (RF) field $\mathbf{B}_1$, tuned to the Larmor frequency, is applied in a direction perpendicular to the main field $\mathbf{B}_0$, it will cause the net magnetization **M** to nutate away from equilibrium orientation along the z-axis towards the transverse plane (Figure 1.5). When the $\mathbf{B}_1$ field is switched off, i.e. at the end of the RF pulse, **M** will continue to precess about the direction of $\mathbf{B}_0$. This precessing magnetization will induce an oscillating signal (known as free induction decay FID signal) in a receiver coil, resonant at ω_0, situated in the transverse plane. This is the signal detected in MR. In a simple model, MR image is a map of the transverse magnetization at a given time by the use of RF excitation pulse [18].

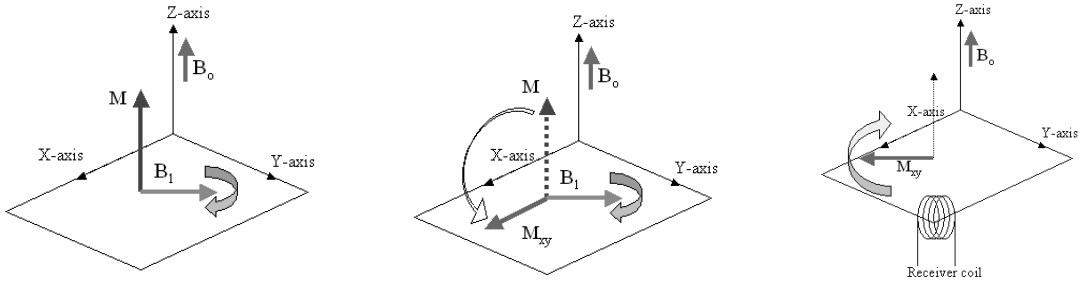


Figure 1.5: Schematic diagrams to show radio frequency field and MR signal received by the coils. (Left) A radio frequency field B_1 rotating at the Larmor frequency (in resonance) produces a transverse magnetization. Note that alignment of the individual magnetization moments creates a net macroscopic magnetization **M** pointing along the main field $\mathbf{B}_0$. (Middle) If the duration of the $\mathbf{B}_1$ field is sufficient to nutate the magnetization **M** by a 90° angle, the entire magnetization ends up in the transverse plane. (Right) Following the radio frequency pulse, the transverse magnetization $\mathbf{M}_{xy}$ rotates about the axis of the main field. An RF signal in the receiver coil situated in the transverse plane will be induced (free induction decay signal) by electromagnetic induction as flux is changing across the coil by the rotational motion of $\mathbf{M}_{xy}$.

In general, the amplitude and duration of the RF excitation pulse can be tailored to nutate **M** through specific angles such as 90° pulse, 180° inversion pulse or 180° refocusing pulse. In reality, the precessing **M** decays gradually via two relaxation processes, namely, spin-lattice and spin-spin relaxation, which are characterized by the relaxation time T_1, and T_2, respectively. Both phenomena cause energy loss from the spin system to the surrounding lattice or exchange of energy between spins. The dynamics of net magnetization can be described by the following two equations for the transverse $\mathbf{M}_{xy}$ and longitudinal $\mathbf{M}_z$ of the macroscopic magnetization at resonance [21],

$$\mathbf{M}_{xy} = \mathbf{A}e^{-\frac{t}{T_2}}$$

$$\mathbf{M}_z = \mathbf{M}_o - \mathbf{C}e^{-\frac{t}{T_1}},$$

where $\mathbf{M}_0$ denotes the full longitudinal magnetization value along the z-axis, **A** and **C** are constants that depend on the initial conditions at the time immediately after the RF pulse excitation. Thus the particular type of pulse sequence employed governs these two constants. It is worth noting that MR contrast depends on the relaxation times T_1 and T_2 in various tissues rather than spin density [22].

To obtain spatial information during image acquisition, magnetic field gradients are superimposed on the static magnetic field. The purpose of these gradients is to specifically relate signal frequency (or phase) to the spatial coordinates in which the net

magnetization of a particular voxel of tissue resides. The field gradient helps to differentiate spatial location of MR signals obtained by the receiver coils. There are different approaches to MR acquisition; however, three steps are generally involved: excitation, phase encoding and readout. In a typical process, the excitation step selects an image slice. A 90° RF pulse is applied to the body in the presence of a field gradient (along z-axis) orthogonal to the slice plane. Within the slice, the RF excitation pulse will selectively nutate **M** from along the z-axis to the transverse plane. It should be noted that the RF pulse has a narrow bandwidth $\Delta\omega$. A thin slice perpendicular to the field gradient of thickness $\Delta z = \Delta\omega / \gamma G_z$ is selectively excited. Because the field gradient can have any direction, a slice of any desired orientation can be imaged. in the phase encoding step, after the slice-selection RF pulse, a gradient G_y is applied causing the nuclei to precess at frequencies which depend on their location with respect to the y-coordinate. At the end of this step, the net magnetization **M** in different locations builds up a phase angle that is uniquely related to the position of the voxels along the y-axis. The readout step follows the removal of the gradient G_y. A readout gradient G_x is applied to create a distribution of frequencies along the x-axis. Depending on the location of the net magnetization **M** along the x-axis, the precessional frequencies of **M** vary and are determined by the field gradient G_x. This precessing magnetization undergoes decay and produces the FID signal in a receiver coil. The phase encoding and the readout steps provide data in the Fourier space. The information gathered from one such cycle (excitation, phase-encoding, readout) is not sufficient to reconstruct an image slice and the cycle has to be repeated with a different setting of the field gradient G_y. For a matrix size of 128×128, the required number of cycles or FID signals is 128, each corresponding to a different value of the phase encoding gradient. Each FID signal detected by the receiver coil is digitized into 128 samples and stored in the memory of the computer. By applying a field gradient for a short time between excitation and signal acquisition, and repeating the process a number of times with different amplitudes of G_y, it is possible to generate a 2D phase encoded MR data in the 2D spatial frequency domain [18,23]. The image can be reconstructed by taking a 2D Fourier transform of the encoded MR data. Alternatively, gradient are set to provide projection data (Figure 1.6) and the image is reconstructed by filtered back projection [24].

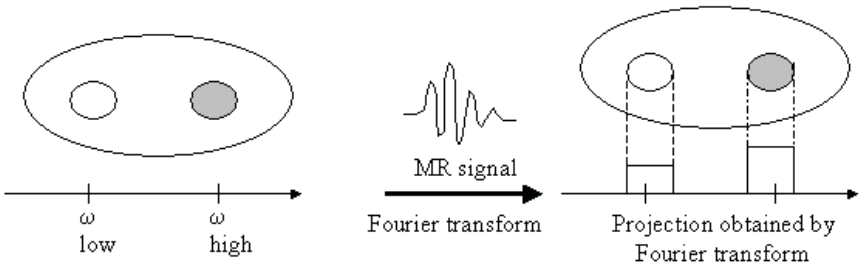


Figure 1.6: If a field gradient G is applied during MR signal acquisition, the resulting signal will be frequency encoded in the direction of the gradient. The Fourier transform of this signal will produce a projection through the object at right angles to the gradient direction. By varying the direction of G, multiple projections are obtained and the image can be reconstructed by filtered back projection.

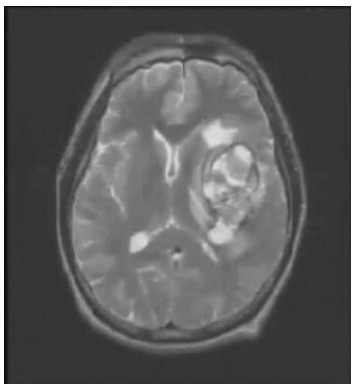


Figure 1.7: A typical MR T_2-weighted image. There is good soft tissue contrast. Note the intense signal from the gray matter around the surface of the brain and from white, grey matter, CSF and skin/epithelium. Abnormal tissue (a large tumour in the left hemisphere) is clearly visible. Note the very weak signal (dark area) is from the bone (skull), which has the lowest water content of all tissues.

Unlike x-ray CT, MR can image slices in any desired plane by an appropriate use of gradients. Many different images may be obtained by varying the RF/gradient pulse sequence, for example, in order to change the T_1 and T_2 weighting. A single slice of a typical image consists of 256×256 or, increasingly, 512×512 voxels. The in-plane resolution is typically 0.2-2.0 mm with slice thickness of 1-5 mm. One of the most important applications of MR is to investigate the brain and spinal cord, where abnormal tissue can be precisely identified (Figure 1.7). Because of its excellent depiction of soft tissues, MR is used to examine suspected tumours, multiple sclerosis [25,26] and degenerative diseases. Cardiac MR and abdominal imaging are areas of increasing importance. A paramagnetic contrast agent [27,28] can be employed to highlight blood flow to investigate, for example, coronary artery disease [29]. MR can also be used to investigate brain function [25,30]. Any brain activity requires energy, resulting in an increased oxygenated blood flow to that area which allows this area to be identified. However, MR is usually employed as an anatomical modality.

1.2.2 Functional imaging based on radionuclide emission

The essential idea of radionuclide imaging is to use a radioactive substance as a tracer to study a particular physiological process in the body. Radionuclide imaging can be used to estimate the function of an organ, monitor chemical transportation in the body or simply examine the spatial distribution of a labeled compound through the body. The ability to follow the movement of radioactive materials through the body was helped by Hal Anger's invention of the gamma camera in 1958 [31]. Further major advances in the early 1970s occurred as image reconstruction algorithms, initially developed for computed x-ray tomography, were applied to nuclear medicine studies and produced transaxial slices of the activity distribution in the body [32]. Two main techniques have been developed, single photon emission tomography (SPET) and positron emission tomography (PET). SPET and PET make use of a wide variety of radioactive isotopes and the chemicals that carry them to particular cells or tissues in the body. For example, radionuclides can be used to assess the efficiency of the heart as a pump, the rate at which the kidneys process blood and eject urine to the bladder or the blood perfusion in brain. These functional studies indicate the presence or absence of disease, and whether or not medication or a surgical intervention is indicated. The efficacy of such intervention or treatment can also be validated by follow-up scans.

Single photon emission tomography

In SPET, radionuclides that emit single photons are used. The typical energy of the photons used for SPET lies in the range of roughly 100 to 400 keV. Gamma photons from radionuclides are detected by a gamma camera taking a series of planar images at successive angles around the patient. Each planar image is analogous to a plain x-ray film. Figure 1.8 shows a single head gamma camera acquiring data from a thorax phantom and the corresponding projective planar image (view). If many views are taken from different angles around the patient, it is possible to reconstruct the tracer distribution in 3D [33]. A typical image data set contains multiple slices, each slice consisting of 64×64 or, increasingly, 128×128 pixels. The typical slice thickness is 4-8 mm. These images demonstrate the concentration of the radiopharmaceutical taken up or excreted by various organs, typically over a time span of several minutes. The image data can be used as a diagnostic indicator of the patient's well-being. The ideal radiopharmaceutical must be taken up rapidly, completely and specifically by only the target organ or tissue under investigation. Therefore, the pharmaceutical used to image

blood flow in the heart will be different to that used to image the brain or kidneys, but both may be labeled with the same radioactive marker.

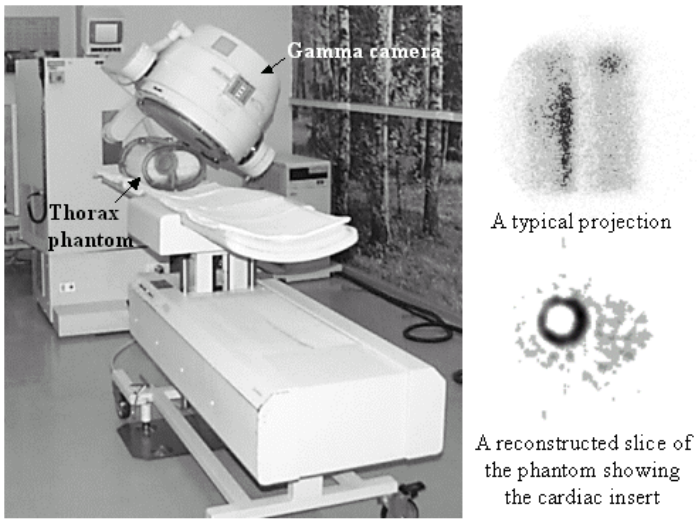


Figure 1.8: A single-head SPET camera is used to scan a thorax phantom (left). A projection acquired by the camera (top right) and a transverse slice of the reconstructed image with the torso and cardiac insert (bright ring in the bottom right image) is shown. The projection and the reconstructed image is shown in reversed colour (Image courtesy of Westmead Hospital)

The radioactive markers most commonly used are technetium-99m (^{99m}Tc) with gamma photon energy of 140 keV, and iodine-123 with gamma photon energy of 160 keV. Various phosphate compounds can be tagged with ^{99m}Tc to provide an image of the skeleton for studying the function of the bone. Red blood cells can be labeled with ^{99m}Tc, providing opportunities for monitoring cardiac function. Gallium-67 citrate uptake provides information on lymphoma or infection. SPET is a useful tool for the management of patients with stroke, epilepsy, recurrent brain neoplasms, and some forms of dementia [34]. Classic problems in SPET involve improving sensitivity, and correcting for attenuation, scattered radiation and distance-dependent resolution [35,36,37,38]. A more detailed discussion of SPET acquisition and reconstruction will be presented in chapter 5. A typical Gallium-67 SPET uptake to detect liver hepatoblastoma and its corresponding MR image is shown in the Figure 1.9. The image illustrates the difficulty in delineating the “functional area” accurately because of poor spatial resolution and low uptake in some tissues of the target region.

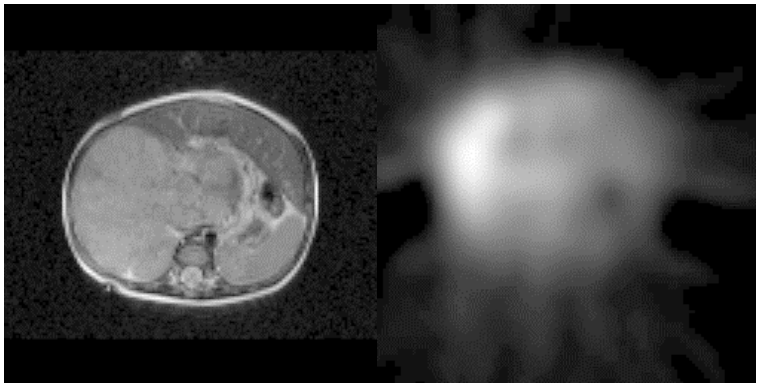


Figure 1.9: A transverse slice of a MR (left) and Gallium-67 SPET (right) in the abdominal region is shown. Note the poor spatial resolution of SPET compared to MR.

Positron emission tomography

In PET, the radionuclide used for labeling is a positron emitter rather than a gamma emitter [39,40]. The emitted positrons combine with electrons to annihilate each other, resulting in the emission of two gamma photons, each of energy close to 511 keV, that travel in opposite directions (Figure 1.10). Since each event results in two photons traveling in opposite directions, the detection of these photons in coincidence on two opposing detectors identifies the line containing the location of the annihilation event, without the need for conventional collimation (Figure 1.11). The data can be reordered into projections and reconstructed in the same way as SPET data [41,42,43]. Examples of positron emitters include oxygen-15 and fluorine-18. The most widely used PET tracer is ^{18}F-fluoro-deoxyglucose (FDG) (Figure 1.12). This substance follows a part of the glucose metabolic chain but then is trapped so that the activity distribution reflects the metabolic rate of glucose. Since most tumours have an elevated glucose metabolic rate, FDG is useful in localising both primary tumours and metastases [39,40]. PET is also used to diagnose heart disease, and it is often used to investigate the function of the brain, providing information about dementia, stroke and epilepsy [44, 45].

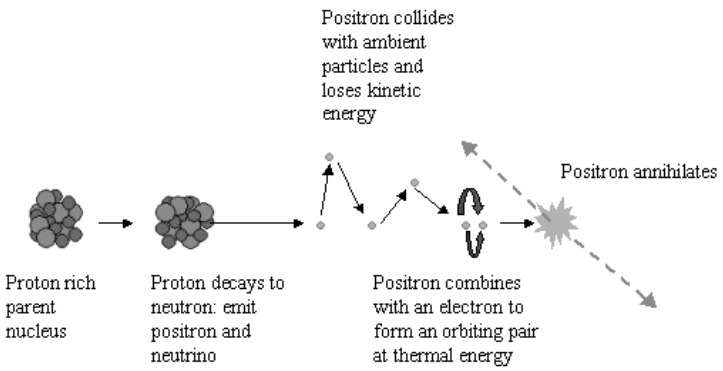


Figure 1.10: Positron emission and annihilation produces two anti-parallel 511 keV photons.

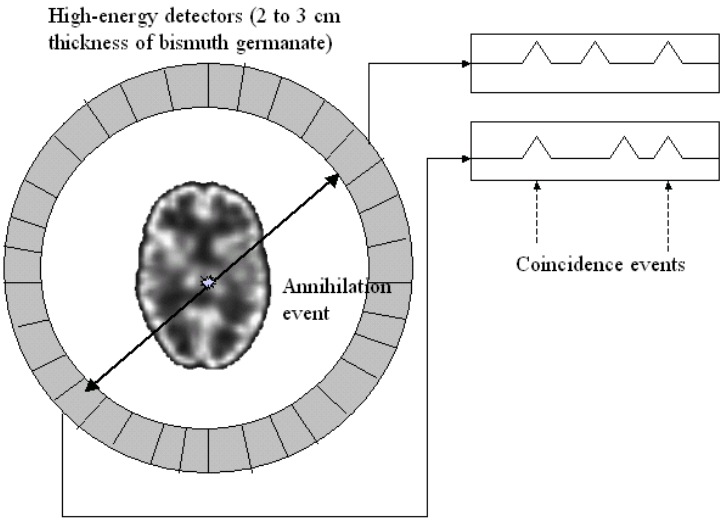


Figure 1.11: The detector consists of a series of rings of detector elements that register a count whenever a photon falls on it. If pulses from opposing detector elements coincide, it is assumed that the registered photons arose from the same annihilation event.

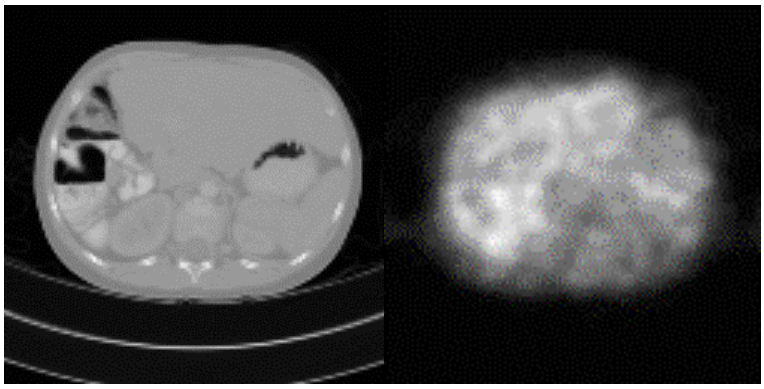


Figure 1.12: A typical CT (left) and ^{18}F-fluoro-deoxyglucose (FDG) PET (right) scan of a transverse slice through the abdominal region of a body. (Images courtesy of Children Hospital, Westmead)

1.3 Why medical image registration?

At the present stage, clinicians can employ a vast array of specialized state-of-the-art instruments to acquire images for the use of medical diagnosis. Some of these image modalities include computed tomography (CT), magnetic resonance imaging (MR), positron emission tomography (PET), single photon emission tomography (SPET), functional magnetic resonance imaging (fMR) and ultrasound imaging (US). Different imaging techniques rely on different physical principles producing images of diverse contrast, signal-to-noise ratio and resolution. Some images (e.g. CT, MR) are intrinsically able to yield high resolution anatomical details and have less noise relative to other images (e.g. SPET, PET). The need for fusion comes primarily from the fact that images from different modalities will provide better diagnosis, management and treatment of patients. For example, PET and SPET provide mainly functional

information and lack anatomic details, and vice versa for CT and MR. Here image fusion helps to associate these functional activity to specific anatomical region. Image fusion is also valuable for longitudinal study. It helps to monitor the growth or change in an organ for certain pathology when images are acquired for the same body region at different time instances. In certain circumstances, information derived from different images is utilized to provide navigational aids for image-guided surgery [46]. In general the integration of images from two different scans is not trivial since the relationship among them and to the scene coordinate system are different. These give impetus to the needs of image fusion, which drive the development of registration to represent information pertaining to the same object system of study in a common coordinate system.

Since medical image fusion has emerged as a particularly active field, there has been much interest in the development of image registration [47,48,49,50]. This has resulted in applications of registration techniques, which help to detect, locate, monitor and measure pathological and other physical changes in patients. The registration of functional (SPET, PET) and anatomical (MR, CT) images permits correlation of changes in metabolism, blood flow, and other functional measurements with regional anatomy and morphological changes, and is important for diagnosis, surgery and therapy [49,51,52]. Due to the high sensitivity and specificity in detecting cancer comparing to CT and MR, PET and SPET scans can distinguish malignant from benign lesions and can identify areas of cancerous involvement. Utilisation of high resolution anatomic images to better localise identified lesions has a significant impact on patient management [53,54]. The registered images can be used to aid image-guided treatments and monitor progress pre-, intra-, and post-therapy [55]. An introduction to the terminology of image registration, classification of registration techniques and a survey of clinical applications can be found in [47,49,50,56]. Some reported clinical applications are given below briefly.

By far the most widely used applications of image registration are in the brain. One particular example is the ictal to inter-ictal ^{99m}Tc-HMPAO SPECT image registration which can be used to define the seizure focus in the diagnosis of focal epilepsy [57, 58,59]. With the incorporation of the subtracted ictal and interictal images, image-guided surgery has been developed to resect these epileptic foci with reference to the

registered MR [60]. Other examples include the studies to detect brain functional changes in Alzheimer's disease compared to normal data [61,62,63,64]. In brain tumors, PET-MR registration is clinically useful for disease staging and therapy control. It allows differentiation of tumor recurrence from scar-tissue, fibrosis and necrosis that may result from therapy [65,66]. Others have suggested using CT-MR brain registration to plan skull base neurosurgery, staging nasopharyngeal carcinoma and radiation therapy planning [67,68,69,70,71,72].

A growing number of papers focus on registration outside the head. Mutic *et al.* [73] suggest the use of multimodality image registration as quality assurance for conformal 3D treatment planning in the neck as well as in the head. Other applications include image fusion of xenon-133 SPET ventilation studies and chest CT images to assess the anatomic extent and location of impaired ventilation. Munley *et al.* [74] reported that pre-radiotherapy SPET and PET refine treatment planning when functional data are available to assist in radiation field design. Several papers demonstrate the use of registration to improve understanding of specific pathophysiology. Somer *et al.* [75] used PET-MR image registration to study soft tissue sarcoma in the knee, thigh, groin, flank or back. Forster *et al.* [76] registered CT and SPET to study neuroendocrine abdominal tumours with liver metastases. Other authors registered PET and CT images of the thorax [77,78], aligned SPET with CT scans in patients with metastatic thyroid carcinoma to identify patients with suspected recurrence of thyroid cancer [79,80], or matched dual-isotope SPET with CT images of endocrine carcinomas in the thoracic-abdominal region [81]. A number of research groups applied a cardiac perfusion image (SPET or PET) registration with a normal template to determine the abnormalities in the heart [64,82,83,84]. Fei *et al.* [55] used registration to assess the feasibility for interventional MR-guided treatment of prostate cancer.

Examples of image registration are also found in other image processing applications, such as partial volume correction, image reconstruction and motion correction [49]. Due to their coarse resolution, functional images manifest significant partial volume effects. There is an inter-regional spillover of counts into adjacent structures, which leads to a reduction of the true activity signal within identified regions [85]. Anatomical MR images can be used to correct this partial volume effect. One approach is to use a segmented MR image to simulate a PET or SPET image. After generating a simulated

"noise free" image from MR and registering it to the measured functional image data, a multiplicative correction of intensity can be performed [85,86,87]. Another use of registration is in the correction for heart motion. This motion induces blurring during PET and SPET studies. Correction of myocardial studies has been reported to reduce artefactual perfusion defects [88,89,90]. Registration has been applied to correct for head motion to minimize motion artefacts in the reconstruction. One approach is to reorient the reconstructed image in 3D until the alignment of the constructed and measured projection is optimal [91,92]. The head motion correction can be extended to image-guided surgical procedures that rely on continuous on-line tracking of the patient's head during surgery [93]. Image registration can also be used to refine SPET resolution using Bayesian reconstruction, where a registered MR study serves as a prior. Smoothing in the reconstruction is constrained to within anatomical boundaries of the MR prior with the aim of improving the reconstructed resolution [94,95,96]. The list of applications is by no means complete. Interested readers are referred to [49].

1.4 Problem statement

To combine different scans of the patient is a non-trivial task because image information cannot be integrated directly. The patient cannot be relied on to assume exactly the same pose at different times, and possibly in different scanners. Furthermore, images can be acquired for the same patients at different time instance in longitudinal study. The aim of image registration, also known as co-registration, alignment and matching, is to calculate the best geometric transformation between corresponding images. Multimodality image registration refers to spatial matching of two images from different imaging modalities (Figure 1.13). The main purpose of this thesis is the registration of 3D anatomical and functional images. When correctly registered, each location in the functional image will reflect physiological information about the corresponding location in the anatomical image. In general, image registration involves rotating, translating, scaling and possibly warping of image grids. Beside the misalignment between scans, there are specific difficulties attributed to the characteristics of functional images. One challenge is the large difference in image resolution between functional and anatomical modalities. PET and SPET are usually much more noisy due to the relatively small photon flux. Furthermore, these images suffer modality specific image degradation due to the limited resolution (e.g. partial volume effect) and artefacts because of motion (e.g. respiration and cardiac motion).

Another challenge is the difference of intensity distribution between images. The ideal radioactive tracer for nuclear medicine targets only one particular type of tumour or tissue with little or no uptake in other tissues. Such local specific distributions are not well suited to registration because of insufficient information to constrain the alignment of images. A further major challenge is that different organs are misaligned in different ways. For example, rotational misalignment in spine between images was reported in [81], and upward and outward misregistration was found in the thoracic region by [97]. Image registration may require non-rigid transformation to align images because of local misalignment of internal organs due to the change in body posture or changes in the state of internal organs (e.g. empty versus full bladder). Thus a robust and accurate registration procedure is necessary to tackle these challenges.

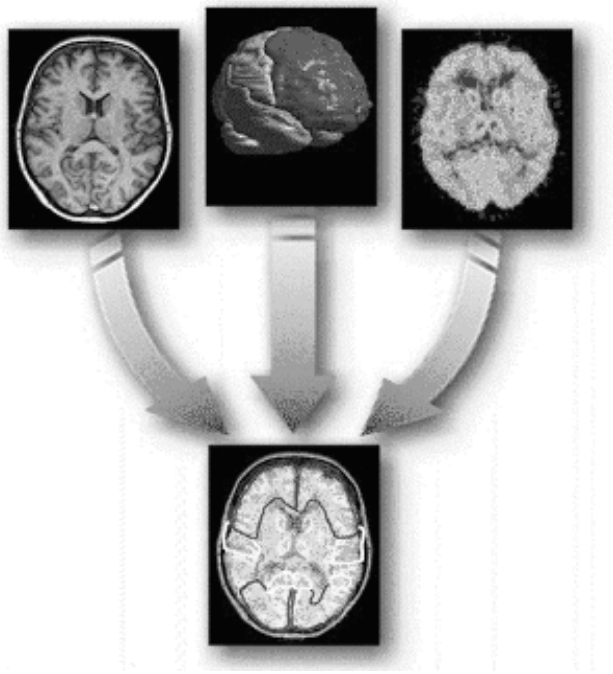


Figure 1.13: Registration of a functional and an anatomical image enables image fusion. Additionally, registration with a brain atlas allows boundaries of brain regions to be delineated on the clinical data. (Image courtesy of Dr. L. Thurfjell)

1.5 Thesis overview

This thesis describes a study of the application of joint intensity algorithms to the registration of emission tomography and anatomical images. The emission tomography images are acquired by targeting appropriate radiopharmaceuticals in SPET or PET to extract specific functional information. CT and MR provide anatomical images. In this thesis, work on rigid body medical image registration is described first, followed by non-rigid registration.

A review of medical image registration literature is given in Chapter 2. A generic structure of the registration process is invoked, incorporating coordinate transformation, intensity interpolation, similarity measures and optimization, in order to describe and compare different registration methods. The second part of Chapter 2 gives an overview of different validation techniques that can be used to assess the accuracy, robustness and precision of a registration algorithm. A theoretical discussion of joint intensity histogram techniques, including entropy-based and variance-based approaches can be found in Chapter 3. Based on the assertion that an image can be taken as a random variable, mutual information, normalized mutual information and correlation ratio are developed. As an extension of the correlation ratio, a symmetrical correlation ratio is proposed. The mutual information and normalized mutual information techniques use entropy as a means to characterize dispersion of intensity distribution in the joint histogram. The correlation ratio and symmetrical correlation ratio are variance-based methods that rely on minimizing the variability of intensity distribution of one image given the observation of intensity in the other image.

In Chapter 4, the symmetrical correlation ratio is compared to other joint histogram techniques in the context of registration of medical images constrained to rigid body transformations. A quantitative evaluation is included in this chapter. Because synthetic misalignment can only simulate real misregistration in a limited sense, clinical data of real misalignment obtained from 12 patients using external fiducial markers are also used for the evaluation of registration accuracy.

Chapter 5 describes the application of a torso phantom to simulate the specific degradation, artefacts and noise in SPET. The segmented digital phantom simulates a distribution of a particular radioactive marker in a clinical SPET image. Furthermore, non-uniform attenuation and depth dependent resolution are incorporated in the simulated SPET. Two popular reconstructions, filter-back projection and accelerated statistical reconstruction [98], are used to simulate various SPET studies from the digital phantom to validate registration methods.

Chapter 6 is an extension of Chapter 5 in that the digital phantom is used as a tool to analyse registration issues in thoracic-abdominal region. Several aspects, which potentially influence accuracy, are controlled in the registration process to investigate

their effect on the performance of registration. These aspects include the choice of the reconstruction algorithm, attenuation compensation, image noise, pre-registration processing, and field of view. Several clinical examples are presented to show the feasibility of the proposed approach. The registered images demonstrate improved localization of functional abnormalities in relation to the anatomy.

Chapter 7 gives the main theoretical description of a novel non-rigid registration. Local misalignments between images are recovered by sub-image block registration. This provides a displacement vector field to warp one image to the other image. This chapter also investigates optimum setting of different parameters such that the computational cost is minimized without sacrificing accuracy of registration. Since it is computationally too expensive to estimate the displacement vector at each voxel, sparse sampling together with a multi-level approach is used to accelerate the process.

Chapter 8 presents a quantitative validation of intra- and inter-modality registration in brain and non-brain regions. In most cases, "ground truth" is difficult to determine. A realistic simulation based on actual clinical data is used to assess the accuracy of registration. In this chapter, the recovery of a known displacement field is evaluated on a point-by-point basis and the mean difference between the derived and actual displacement fields calculated. The validation study is described for MR-MR T_1 images, MR-T_2-proton density images, and MR-SPET images in the brain region, as well as CT-SPET images in the thoracic-abdominal region.

The final chapter summarizes the findings of the investigations described in this thesis and gives suggestions for further work.

Chapter 2

Review of medical image registration

2.1 Introduction

Medical imaging plays a vital role in clinical diagnosis, as well as planning and evaluation of therapy. Several imaging modalities are in widespread use in current clinical practice. They can be divided broadly into anatomical and functional modalities. Anatomical modalities, such as x-ray computed tomography (CT), magnetic resonance (MR) imaging and ultrasound imaging, show detailed anatomical structure. Functional modalities, including single photon emission tomography (SPET) and positron emission tomography (PET) as well as functional MR (fMR), depict the physiological and pathological status of tissues or organs. The appearance of some of these images can be seen in Figure 2.1. Images gained from anatomical and functional modalities can provide complementary information for diagnosis, radiation therapy planning, surgical guidance, and monitoring of treatment. Integration of these images is often desired. For example, in radiotherapy treatment planning, a CT scan is needed for dose distribution calculations while the contours of the target lesion are often best outlined on MR images. In nuclear medicine, localization of dysfunctional areas is helpful in planning surgery for focal epilepsy. A critical step in this process is the spatial alignment of images.

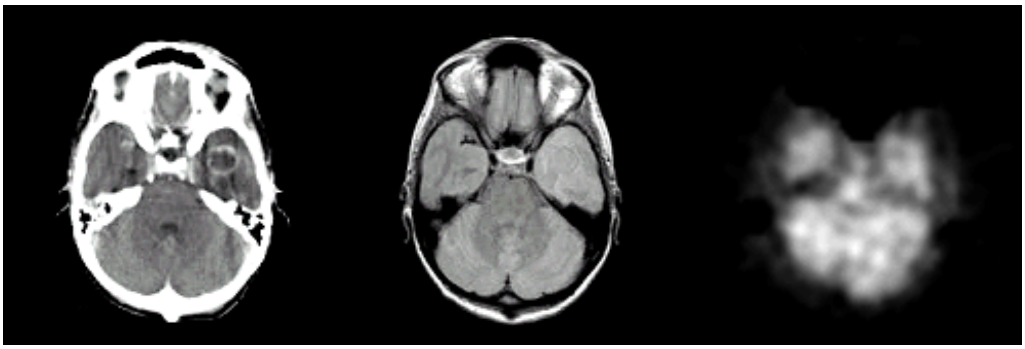


Figure 2.1: Examples of clinical images showing transverse sections of the brain. CT, MR and SPET. (from left to right)

There are numerous ways to categorize image registration methods. Brown [99] surveyed registration in the context of remote sensing, computer vision and medical imaging. Van den Elsen *et al.* [56] compiled an overview of medical image registration and, more recently, Maintz and Viergever [47] comprehensively surveyed and

systematically classified medical image registration methods. Lester and Arridge [100] discussed elastic image registration in detail. Hill *et al.* [48] reviewed registration algorithms and provided an account of the development of voxel-based registration. A number of recent reviews focused on particular areas of application. Hawkes [101] focused on registration of anatomical images in diagnostic radiology and their clinical application. Hutton *et al.* [49,102] reviewed image registration and its applications in the context of nuclear medicine. More recently, Pluim *et al*. [50] gave a comprehensive survey of image registration based on mutual information as an introduction to those new to the field.

2.1.1 Classes of application

Applications of medical image registration can be classified according to the subject and the image modality. Intrasubject registration is necessary for comparison of serial studies of the same patient, but intersubject registration can also be of interest, particularly in medical research. The images can be of the same kind (intramodality) or acquired by different techniques (intermodality). Following Maintz and Viergever [47] and Hutton *et al.* [49], four basic classes of application are distinguished.

Intra-subject, intra-modality applications include treatment verification by comparison of pre- and post-intervention images, comparison of ictal and inter-ictal SPET images in epileptic patients, time series MR or CT scans to monitor growth of tumours, and fMR time sequences. Registration of CT (x-ray attenuation) and MR (tumour delineation) will benefit radiotherapy treatment planning. Another example is the registration of a functional study (SPET or PET) with an anatomical scan (MR or CT) to align anatomical and functional data. This allows localization of increased accumulation of a radiopharmaceutical within specific anatomic structures. Examples of intersubject, intramodality applications include measurements of cross-population variability [103,104] using statistical parametric mapping (SPM) [105,106,107,108], and matching of an individual subject's image to an anatomical reference atlas [109]. Finally, alignment of images of different subjects across different modalities (intersubject, intermodality) is required in extracting group characteristics of neuronal function, as related to an anatomical framework.

2.1.2 Dimensionality

The domain of the image is typically 3D (CT, MR, SPECT, PET) but some modalities are intrinsically 2D (x-ray projection imaging, planar radionuclide scans, conventional ultrasound imaging). However, anatomical structures are intrinsically 3D so that even problems posed as 2D-2D may require determination of 3D transformation for successful registration. Registration in serial digital angiography is a 2D-2D image alignment. With this technique, a sequence of 2D digital x-ray projection images is acquired to show the passage of injected contrast material through the vessels of interest. Background structures are largely removed by subtracting an image acquired prior to injection. The subtraction is made ineffective by patient motion. To correct for the motion, a 2D geometrical transformation that accounts for the projective effects of a 3D transformation is required. The registration is difficult since the projection geometry is unknown. Interested readers may refer to Meijering *et al.* [110] for a comprehensive review of motion correction in digital subtraction angiography.

The 2D-3D registration has been applied to correct for patient motion in SPET reconstruction [91,92,111]. These algorithms seek a projective transformation of the 3D image that comes closest to the observed projection data. Another example is the registration of a portal image to a CT image to verify patient positioning for radiotherapy [112,113,114]. Penney *et al.* [115] compared algorithms that register a 2D fluoroscope image to a CT volume. This thesis is primarily concerned with 3D-3D image registration. In most instances, the term ”image” will refer to a full 3D image data set.

2.1.3 Generic algorithm for image registration

The aim of image registration is to determine a spatial transformation that aligns an (floating) image to a second (reference) image to obtain the best correspondence. A schematic diagram in Figure 2.2 is used to illustrate some major steps for rigid-body images registration. Points in xyz space can be expressed as vectors of homogeneous coordinates where they are represented as a four-element column vector of three coordinates, for example: $x \equiv (x_1, x_2, x_3, 1)^T$. One advantage of this approach is that translation, which normally must be expressed as an addition, can be represented as a matrix multiplication. These position vectors are translated, rotated, scaled, and skewed by multiplying them by rigid-body transformation matrices. Thus, to express a

transformation that includes translations as a single transformation matrix **Q**, the position vector will be expressed in homogeneous coordinates $(x_1, x_2, x_3, 1)^T$ and **Q** will be a 4×4 matrix. Let a rigid-body transformation T be described by a mapping of image A to A′, $T_\theta = A \rightarrow A'$, where the transformation is parameterized by θ. In general, this can comprise both spatial and intensity changes but here only global geometrical transformations, $A'(x) = T_\theta A(x) = A(\mathbf{Q}\,x)$ are considered.

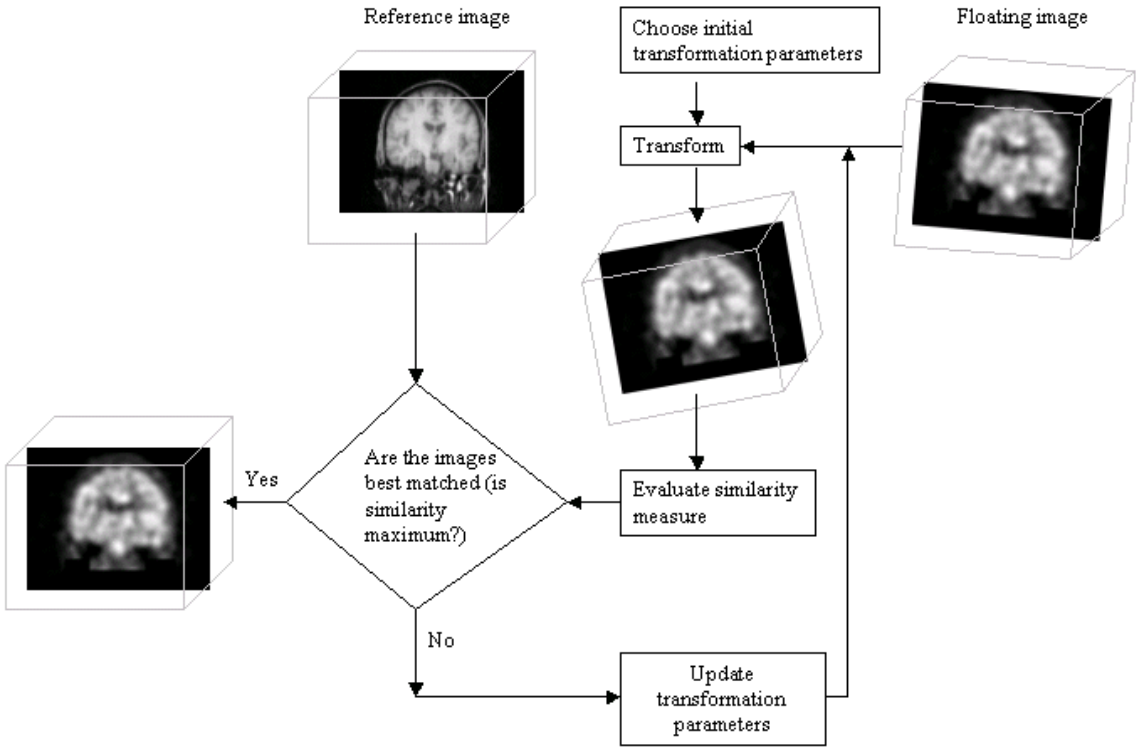


Figure 2.2: A general scheme for image registration. At each cycle of iteration, a measure of similarity between the transformed 3D floating image and the 3D reference image is computed until a global maximum is reached.

Registration algorithms aim to find a transformation that maximizes similarity between the transformed floating image $T_\theta A$ and the reference image B. Let Φ denotes the similarity measure. Then the outcome of the registration algorithm is given by $\hat{T}_\theta = \arg\max_T \Phi(T_\theta A, B)$. The many reported image registration algorithms have certain features in common. In general, four main components can be identified: the similarity measure, the spatial transformation, interpolation, and optimization. The similarity measure determines the "goodness" of alignment between images. The transformation performs the spatial alignment. Interpolation fills in the image values at required points. Optimization usually follows a search strategy in a least-squares [116]

fitting of two 3D points sets to determine the transformation for the maximum similarity between images. When automatic or semi-automatic registration algorithms use a search strategy, they often proceed by iteratively adjusting the spatial transformation parameters until a global similarity measure computed between the transformed floating image and the corresponding reference image reaches maximum or another stopping criterion is satisfied. A detailed discussion of the various elements of the registration algorithm follows.

2.2 Similarity measures

All automatic and semi-automatic registration algorithms require an explicit similarity measure to quantify the alignment between the two images. When the two images are *optimally* registered over the overlapping volume, the similarity measure is expected to reach a maximum value. In clinical applications, the *true* registration is unknown and it is expected that the registration results will depend on the image modality, image noise, distortion (global or local), and the body part (brain or non-brain). Other factors such as the dimension of the parametric space, as well as the behavior and capture range of the similarity measure, also affect the accuracy and robustness of registration. Implementation aspects of the algorithm, including the choice of techniques for transformation, interpolation and optimization, will also affect the performance of registration. Current registration algorithms can be divided broadly into two approaches: feature-based and voxel-based. All similarity measures can be calculated globally (on the entire image) or locally (on a subimage). The choice between these two approaches is based on image characteristics as well as the context of the application.

2.2.1 Feature-based measures

It is sometimes possible to extract suitable features from the two images to be registered. These extracted features can be points, edges or surfaces, yielding progressively more information from the image data for the registration process. Because these features contain relatively few data points, registration is, in general, faster than the voxel-based approach. The similarity measure would typically be formulated as the negative of the Euclidean distance between the sets of corresponding features,

$$\Phi(\mathrm{T}_\theta \mathrm{A}, \mathrm{B}) = -\sum_{\mathrm{i}=1}^{N} || \mathrm{A}(\mathbf{Q}\, x_\mathrm{i}) - \mathrm{B}(x_\mathrm{i}) ||^2 , \qquad (2.1)$$

for a set of *N* feature points x in image A and B. Maximizing the similarity measure is equivalent to minimizing the distance between the corresponding feature sets. If noise (white and Gaussian) is additive, the maximum of this measure corresponds to the maximum likelihood solution [117,118]. The main drawback of this measure is its lack of robustness in the presence of severe outliers. Outliers have to be rejected otherwise the maximum may become less pronounced [119].

Point-based measures

Point measures are used if a small number of corresponding pairs of points are available in the two images. These points, also known as markers, can be extrinsic, anatomical or geometrical. Extrinsic markers are typically fiducial markers either on the patient's skin or attached to stereotactic frames [120,121]. The markers should be visible in both the floating image A and the reference image B. Their form and composition depends on the modality (e.g. $CuSO_4$, $NiCl_2$ or petroleum jelly for MR, iodinated compounds for CT, or gamma emitters for SPET). They vary in terms of ease of implementation and patient tolerance. The most invasive is the stereotactic frame which has to be screwed into the skull [121]. The location error of the markers needs to be accounted for; e.g. the movement of skin between scans is a problem with skin markers. The most important limitation is that extrinsic markers require patient preparation and are therefore unsuitable for retrospective image registration. External markers also have been used to validate registration algorithms [122,123]. In this case, the residual displacement of the surface markers following registration indicates the performance of a registration. In this case, registration is considered optimal when the mean Euclidean distance between corresponding sets of points is minimum. In the case of rigid body transformation, Equation (2.1) can be treated as a least-squares fitting of two 3D sets of points known as the Procrustes problem [48,116,124,125,126]. Given two sets of *N* non-coplanar points $\{x_i\}$ and $\{y_i\}$ where $i = \{1,2,...,N\}$, the negative Euclidean distance Φ to be maximized is given by $\Phi = -\sum_{i=1}^{N} || \mathbf{Q}(x_i) - y_i ||^2$. Where $\mathbf{Q}$ is a 4×4 matrix representing a rigid-body transformation, the problem can further be simplified by referring the coordinates to the respective centroids of each point set. Let $y_i^{'} = y_i - \bar{y}$, $x_i^{'} = x_i - \bar{x}$ and $\mathbf{Q}'(x_i) = \mathbf{R}' x_i + \mathbf{t}$, where $\bar{y}$ is the centroid of $\{y_i\}$, $\bar{x}$ is the centroid of $\{x_i\}$, $\mathbf{Q}'$ is a 3×3 rigid-body transformation matrix, $\mathbf{R}'$ is a 3×3 orthonormal rotation matrix and $\mathbf{t}$ is a 3-element

translation vector, the original least-squares problem is further reduced to $\Phi = -\sum_{i=1}^{N} || \mathbf{R}' x_i' - y_i' ||^2$. The optimum transformation $\hat{T}$ is decoupled and estimated in two steps: First, rotation $\hat{\mathbf{R}}$ that maximizes Φ is found by singular-value decomposition (SVD) [127] of the matrix $\sum_{i}^{N} x_i' y_i'^{T} = \mathbf{UDV}^T$, where the superscript ^{T} denotes transpose, $\mathbf{D}$ is a diagonal matrix and $\mathbf{U}$ and $\mathbf{V}$ are orthonormal matrices. The rotation $\mathbf{R}'$ is given by $\mathbf{R}' = \mathbf{VU}^T$. Second, the translation $\hat{\mathbf{t}}$ is given by $\hat{\mathbf{t}} = \bar{y} - \hat{\mathbf{R}}\,\bar{x}$. Interested readers may refer to the work of Arun *et al.* [116].

External skin markers provide a measurement of the registration error at the body surface and results may not extrapolate to the points inside the body. In many cases, since the definition of the measure does not involve averaging, and the number of markers F are much less than the number of voxels N, then $\sum_{i=1}^{F} \left\| \mathbf{Q}\, x_i - y_i \right\|^2 < \sum_{j=1}^{N} \left\| \mathbf{Q}\, x_j - y_j \right\|^2$. In this case the similarity measure Φ does not provide an upper bound. However, when the external markers are attached to a stereotactic frame, the greater the distance of the markers from the image centre (assume that the rotation is about the image centre) the greater is the rotational displacement of the markers. Thus they may provide an upper bound of average residual registration error within the whole image. Several groups suggested various anatomical landmarks for registration of brain images [120,128]. These can be any localized prominent features in bones, soft tissue organs or blood-vessels. Localization of features in both images is generally carried out by an expert observer and is labour intensive. Based on external fiducial markers, inter-modality registration of the cervical spine [129] and of the brain [121] was performed. Due to the intrinsically poor resolution, anatomical landmarks are difficult to identify in SPET and PET. Such difficulty is partly overcome with a dual-isotope SPET image in which a tumour scan and a bone scan are acquired simultaneously. Internal landmarks in the anatomical image and the SPET bone image were used to determine the transformation that registers the anatomical image and the tumor SPET image [81,130]. This method is only applicable to some specific cases and it increases the radiation dose to the patient. Like the external fiducial markers, anatomical markers can be used to assess the accuracy of inter-subject brain registration

[128,131] or to validate registration algorithms in terms of the target registration error [123,125].

Geometrical landmarks are extremal points (such as corners), which can be extracted by a differential operator from the line of maximum curvature or as a singularity of the isointensity surface in 3D images [132]. In a high resolution CT or MR image, the extremal points method is reliable and works automatically. In SPET and PET images, which are characterized by low resolution and low signal-to-noise ratio, the object surface is generally not sufficiently well defined for the extremal points methods.

Curved-based or geometrical feature-based measures

Edges derived from intrinsic structures such as object boundaries in 2D or curvatures in 3D isointensity surfaces (e.g. skull, brain) can be used to register images when these structures are delineated in both images. The geometrical features defined by isointensity surfaces often coincide with anatomical (object) boundaries defined by segmentation. Gradient and higher order derivative operators are applied to locate edge features. Differential operators incorporate smoothing (e.g. Gaussian kernel) to reduce the detection of noise-induced edges and sometimes to allow multiscale algorithms to be used. In the 2D curves method, differential operators such as the Laplace operator or the Canny detector [133] can be used to detect edges or ridges (crest lines) in the image. Tissue boundaries can also be extracted by a deformable model-based approach that ensures continuity and smoothness of boundaries, such as the active contour model [134]. The deformable models still use differential operators to measure the strength of the edge but their performance can be improved by incorporating prior knowledge of the shape (e.g. smoothness) of the boundary or characteristics of the region enclosed by the boundary such as the minimum variance criterion [135]. Registration is optimal when the least-squares distance between the corresponding edges is a minimum. Feature curves defined by the intersection of the interhemispheric plane and the skull surface have been used to register MR and PET brain images [136,137,138]. The method relies on intrinsic symmetry property of the brain, an assumption that may not apply in other parts of the body. Maintz *et al.* [139] made use of ridge seeking operators to register CT and MR brain images. 3D feature curves are detected as extrema of curvature in the image isointensity surface. Thirion *et al.* [132] defined the feature curve as the line of maximum curvature (crest line) and used it to register CT images

[140,141]. Gueziec [142,143] used a B-spline surface to build ridge lines and used them for registration of two CT images [119]. Once the corresponding feature curves in the two images were established, an iterative method was used to register brain images [144]. The edge method is constrained by the availability of good quality, high resolution images in which corresponding tissue boundaries exist. However, it cannot be assumed that edge features will be replicated in images of different modalities. For example, in SPET and PET, the grey levels measure local concentration of a radiopharmaceutical administrated to the patient which may not correlate well with regions enclosed by anatomical boundaries. Similarity measures based on edges are unlikely to be useful for registration of functional and anatomical images.

Surface-based measures

There exist a wide variety of registration methods using 3D surface information. A detailed summary was complied by Audette *et al.* [126]. In high resolution CT and MR images, surfaces formed by tissue boundaries are commonly used. Three methods dominate surface registration: the head-hat algorithm, the iterative closest point algorithm and the chamfer matching algorithm. The head-hat algorithm, proposed by Pelizzari *et al.* [145], is one of the better known surface-based methods for the registration of anatomical and functional brain images. Surfaces are formed from 2D contours extracted slice-by-slice from each image. The surface in the high resolution anatomical image serves as the 'head' and a relatively sparse set of points in 3D extracted from the contours in the low resolution functional image serves as the 'hat'. Registration amounts to fitting the hat to the head by minimizing the distance of the hat from the head surface. A similar approach was applied to the registration of CT and PET images of the thorax by Yu *et al.* [78]. They extracted the pleural surface of CT and PET transmission scans to define the head and hat, respectively. In some cases, the brain surfaces delineated in both the MR and the PET image were used [146,147]. Based on the same approach, an external fiduciary band was used to define hat points in the SPET image which were aligned with the head surface obtained from the CT image [79].

In the iterative closest points (ICP) algorithm [148], images are registered using points on a 3D surface (e.g. surface curvature values at extrema) extracted from the image. The ICP algorithm searches, iteratively, for the transformation that minimizes the mean-

square distance from points on the floating image surface to the closest points on the reference image surface. The algorithm then determines a new closest point set (surface points) and continues until the mean-square distance reaches a threshold or stabilizes to within a set tolerance. Corresponding point pairs are used to register the images by using either a least-squares approach or a modified algorithm [149]. The modified ICP approach has been used for the registration of high resolution intramodality images [150,151,152]. Declerck *et al.* [83] reported an application of the ICP algorithm to the registration of cardiac SPET images.

In the chamfer matching algorithm [153,154], the feature surface forms a boundary of a segment in the reference image. For each image voxel, the algorithm finds its closest point on that surface and the distance to it. Thereafter, all voxels in the image volume have a value corresponding to the distance to the nearest surface pixel. The distance map does not provide information explicitly about the correspondence between point pairs on the corresponding surfaces of images. In this case, the distance map is used as a potential function to compute a distance potential of a floating image with respect to the surface points in the reference image. Registration is optimal when the total distance potential is a minimum (maximum similarity). The chamfer matching algorithm has been applied in the CT to MR brain image registration [68,155,156]. Based on the strong signal from the skull in CT and the corresponding weak signal in MR, van den Elsen *et al.* [157] extracted the feature curves of the skull as ridge-lines in CT and trough-lines in MR. The chamfer method was then applied to register the CT and MR brain images. Applications of the chamfer matching algorithm include registration of MR-PET and MR-SPET brain images [158,159] and CT-PET lung images [160] limitation of this algorithm is the uncertainty in determining the skull outline in the SPET and PET images.

2.2.2 Voxel-based measures

The feature-based measures require that the intensity values be processed with a detector to extract the corresponding features in each of the images separately. But feature detection generally requires good signal-to-noise ratio, resolution and tissue contrast in image, and may not work well in functional images. By contrast, voxel-based measures are robust in that, in general, they are derived from image statistics and do not require feature detection. Voxel-based measures include intensity difference,

cross-correlation, minimum variance, and measures related to the joint intensity histogram. Measures based on moments or principal axes are sensitive to differences in shape and intensities between images [161,162,163]. They are of limited use in registering functional and anatomical images and therefore are not discussed further in this review.

Intensity-difference based method

This is an intuitive measure that computes, voxel by voxel, the sum of the differences of intensity values between the floating image and the reference image. Based on the intensity difference, the distance measure Φ in Equation (2.1) has two versions as expressed in Equation (2.2) sum of the absolute differences (SAD) when $n = 1$ and square root of sum of squared intensity differences (SSD) when $n = 2$ which minimizes Φ in a least-squares sense during registration [164,165,166]. The sum is over N voxels in the overlapping image volume,

$$\Phi(\mathrm{T}_\theta \mathrm{A}, \mathrm{B}) = -\left\|\mathrm{T}_\theta \mathrm{A} - \mathrm{B}\right\|_n = -\sqrt[n]{\sum_{i=1}^{N} |\mathrm{A}(\mathbf{Q}\, x_i) - \mathrm{B}(x_i)|^n} \qquad (2.2)$$

The $n = 1$ case corresponds to the L1 norm and $n = 2$ case to the Euclidean L2 norm. The intensity difference measures have been used for intrasubject intramodality registration, such as serial MR registration or ictal and iter-ictal SPET studies [167,168]. They rely on the assumption that intensity values in the two images are similar. Because image intensities obtained from different modalities are intrinsically dissimilar, methods based on intensity differences are problematic for multi-modality registration. Both methods are sensitive to outliers (e.g. where large intensity differences between the images are due to a contrast agent). Eberl *et al.* [165] showed SAD to be more robust than SSD and demonstrated it to be widely applicable in clinical practice.

Cross-correlation based measure

The cross-correlation function measures the statistical relationship between intensity value of images A and B,

$$\Phi = \frac{\sum_{i=1}^{N} [\mathrm{A}(\mathbf{Q}\, x_i) - \overline{\mathrm{A}}][\mathrm{B}(x_i) - \overline{\mathrm{B}}]}{\sqrt{\sum_{i=1}^{N} [\mathrm{A}(\mathbf{Q}\, x_i) - \overline{\mathrm{A}}]^2 \sum_{i=1}^{N} [\mathrm{B}(x_i) - \overline{\mathrm{B}}]^2}}, \qquad (2.3)$$

where $\overline{A}$ and $\overline{B}$ are the mean intensities in the respective images and N is the number of voxels in the overlapping image volume. Compared to the intensity difference measures, the intensity assumption is relaxed allowing the intensity values of the two images to be linearly related. The measure has been applied in intra- and inter-modality registration [169,170,171]. Since a linear intensity relationship cannot be relied on across modalities, a slightly different approach is needed for inter-modality registration. In this case, an intensity mapping is estimated and applied to create a floating image that has a similar intensity distribution to that of the reference image. The images are then registered by cross-correlation. Maintz [139] introduced a morphological operator that can be applied to both MR and SPET brain images to bring out edge information for use in cross-correlation based registration [169].

Minimum variance based measure

For multi-modality registration, a simple intensity cross-correlation function cannot be used as a similarity measure because intensities in different modalities are intrinsically dissimilar. To circumvent this difficulty, minimum variance measures [131,172] have been proposed. They assume each image to be approximately piecewise constant (same tissue manifests similar intensity). When transformation is optimal, a region of uniform intensity in the reference image (MR) will correspond to a region of uniform intensity in the floating image (PET). In order to register the lower resolution image (PET or SPET) to the higher resolution image (MR or CT), the higher resolution image is first segmented. Woods *et al.* [172] used a level-set segmentation based purely on intensity. Ardekani *et al.* [131] used a more general approach in which segmentation was accomplished by K-means classification followed by connected component analysis. For any spatial transformation, the segmentation is "induced" in the lower resolution image. The registration algorithm seeks to maximize the uniformity (minimize intensity variance) of the voxel values within each induced segmented region.

In the method of Woods *et al.* [172], let b denote intensity of the reference MR image B. Each b defines an intensity-labeled segment (a level surface) of B, $\{x \in \Omega_B : B(x) = b\}$ and Ω_B is the image domain of B. If there are M possible intensity values so that $b \in \{0,1,2,...,M-1\}$, image B will be partitioned into M segments, some of which may be null. The level-set of the MR image generates M sparse binary images,

one for each intensity, assuming no local correlation. It is worth noting that the partition splits the image (strictly speaking image domain) into "disjoint" subsets $N_{\mathrm{b}} \in \{N_0, N_1, N_2, ..., N_{M\text{-}1}\}$ whose union is the image. Also, the term "partition" applies to the set of subsets rather than an individual subset in the rest of this Chapter. For registration of the floating image A to the reference image B, a similarity measure is computed over the "induced" partition of the transformed image of A. Let $\mathrm{A}'(x) = \mathrm{T}_\theta \mathrm{A}(x) = \mathrm{A}(\mathrm{Q}x)$ be defined as in the Subsection (2.1.3), by interpolation, at grid points of B (a subset of Ω_B). Only the common part of the grid Ω_{AB} (i.e. $\Omega_{AB} = \Omega_{A'} \cap \Omega_B = \Omega_{TA} \cap \Omega_B$) is used to compute Φ. For Woods' algorithm, the partition is confined to the common grid and defines $S_b = \{x \in \Omega_{AB} : B(x) = b\}$. This also partitions (this is also referred to as induced segmentation) the common part of the floating image A′. The statistics of each induced segment can then be examined in A′: $\mu_b = \frac{1}{N_b} \sum_{x \in S_b} \mathrm{A}'(x)$, and $\sigma_b^2 = \frac{1}{N_{\mathrm{b}}} \sum_{x \in S_b} [\mathrm{A}'(x) - \mu_b]^2$. Woods' similarity measure is defined as

$$\Phi = -\sum_{b=0}^{M-1} \frac{N_b}{N} \frac{\sigma_b}{\mu_b}, \qquad (2.4)$$

where μ_{b} and σ_{b} are the mean and the standard deviation of voxel intensities of the segment in A′ induced by intensity *b* in B.

Ardekani *et al.* [131] took correlation among neighboring voxels into account in segmentation and partitioned the image into *K* connected components S_k , $k \in \{0,1,2,...,K-1\}$. The similarity measure Φ becomes $\Phi = -\sum_{k=0}^{K-1} \sigma_k^2$, where $\sigma_k^2 = \frac{1}{N_{\mathrm{k}}} \sum_{x \in S_k} [\mathrm{A}'(x) - \mu_k]^2$. The statistical parameters μ_{k} and σ_{k} are defined in exactly the same way as in the Equation (2.4) except the subscript *b* is replaced by *k*. In the Woods' approach, *b* stands for intensity and a location index in A′. For Ardekani's work, *k* stands for location index rather than intensity.

Joint intensity histogram based measure

Inspired by the success of Woods' algorithm [172], Hill *et al.* [173] proposed an adaptation of woods' methods by constructing a feature space to capture the relationship between intensities in different images. The feature space is a 2D intensity plot showing the combinations of grey values in each of the images for all corresponding points. Figure 2.3 shows such a feature space plot for an MR image with itself. Compared to the feature-based registration, the main advantage of these approaches is the capability for intra- and inter-modality image registration. The 2D intensity plot is generated by plotting the intensity of the transformed image A′ against the intensity of the corresponding voxel in the reference image B where $a = \mathrm{A}'(x) = \mathrm{A}(\mathbf{Q}\,x)$ and $b = \mathrm{B}(x)$.

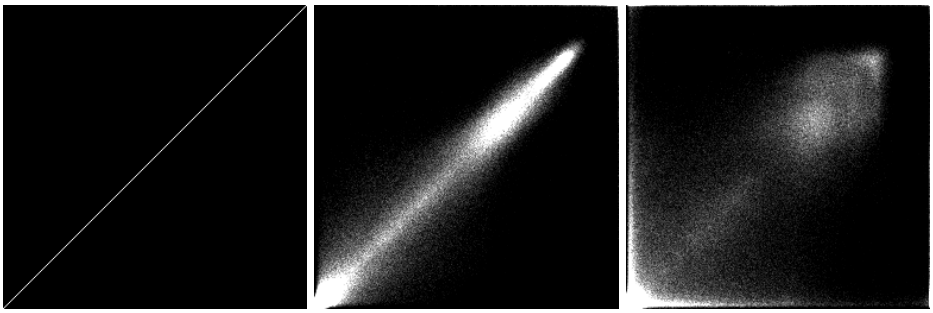


Figure 2.3: A joint intensity histogram created from a pair of images (A and B). The intensity at each point (*a*,*b*) represents the number of voxels in the overlapping region of the two images that have an intensity '*a*' in A and '*b*' in B. A high intensity can be interpreted as a high probability of obtaining the particular pair of intensities in corresponding voxels of the two images. Starting from perfect alignment (left), a relative translation by 1 pixel (middle), and 5 pixels (right) reduces the number of coincidences and disperses the data. To improve display, each histogram was independently scaled to its maximum value.

Compared to [131,172], a region of similar intensity values is defined in the 2D intensity plot instead of in the images. The underlying model for these approaches relies on the uniformity of intensity in one image with respect to the intensity in the second image. Alternatively, a relationship is assumed to exist between corresponding groups of voxels in the images but without any assumption about the form of relationship between voxel intensities. Each entry in the histogram represents the number of voxels with intensity $a \in \{0,1,2,...,M-1\}$ in one image that coincides with voxels of intensity $b \in \{0,1,2,...,M-1\}$ in the second image. Again, M denotes the number of intensity levels in each image with the assumption that both have the same levels. The normalized histogram estimates a joint probability distribution of intensities. Based on the probability distribution, several methods have been derived for the measurement of the dispersion of the histogram. These include joint entropy [174,175], mutual

information [176,177,178,179], normalized mutual information [180] and correlation ratio [181,182]. These measures are constructed in the intensity domain and do not depend explicitly on spatial relationships between voxels. Instead, a spatial relationship of intensity values is implied in the construction of the intensity pair $(a,b) = \left(\mathrm{A}'(x), \mathrm{B}(x)\right)$ at location x. The uniformity of the 2D joint histogram, as measured by entropy or variance, is at a minimum when images are aligned correctly. West *et al.* [183] and Kagadis *et al.* [184] compared the performance of volume-based techniques with the surface-based techniques. They showed that, for CT-MR registration (and, to a lesser extent for the PET-MR registration), the joint intensity histogram techniques give more accurate and more reliable results than the surface-based methods. By far the most common joint intensity histogram measure currently used for anatomical and functional image registration is the mutual information. A detailed discussion of these measures is given in Chapter 3.

2.3 Transformations

2.3.1 Rigid-body transformation

The transformation used in the registration problem can be classified into two main types, namely the rigid and non-rigid models. An illustration of Types of transformations categorized by domain (global/local) and elasticity (rigid-body and nonrigid) are shown in Figure 2.4. The rigid-body transformation preserves angles and lengths so geometric relationships among points in the image do not change. In general, the rigid transformation is applied for intrasubject intramodality registration (e.g. fMR time sequences) or intermodality registration where it is assumed that the distance between any two points in the object is the same regardless of how the object is scanned. That distance must be measured in physical units rather than voxels, as voxel dimensions change from one modality to another.

The validity of the rigid-body model should be considered separately for intra- and intermodality applications. In intramodality applications, it is assumed that the misalignment can be approximated by rigid-body transformations. This implies that the acquisition protocols and imaging systems are the same or, more realistically, that the differences in acquisition can be approximated by rigid-body transformations. The approximation may fail in (e.g. MR) due to nonlinear field distortions. In the same

subject, the rigid-body model is a reasonable assumption for at least some parts of the anatomy (bones and bone-encased volumes, such as the brain), unless they are nonrigidly distorted by a pathological condition. In other parts of the body, the approximation is not as good due to nonlinear distortions arising from posture changes. For inter-subject applications, the rigid-body assumption is not warranted. In the intermodality intra-subject applications, a body part, such as the brain, can usually be assumed to be a rigid body irrespective of the modality. In addition, the rigid-body assumption must also capture differences between modalities.

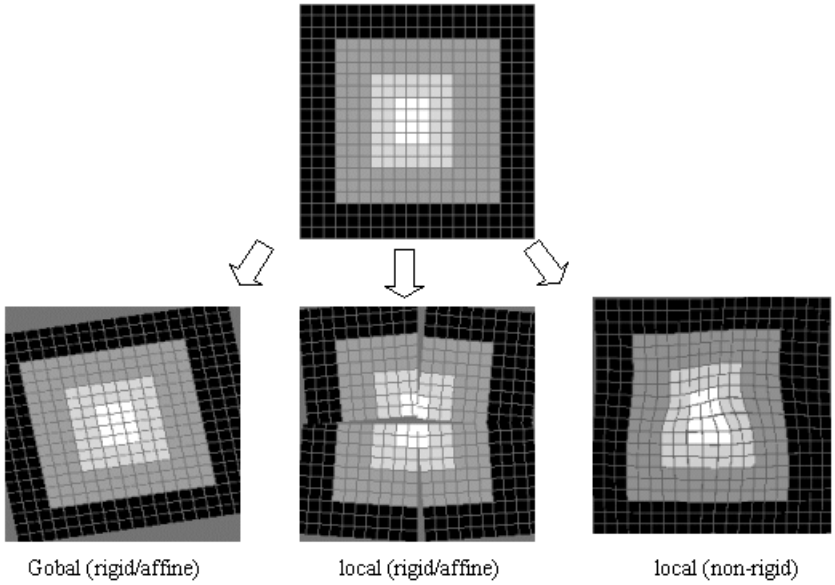


Figure 2.4: Types of transformations categorized by domain (global and local) and elasticity (affine and non-affine). Local affine transformation may induce discontinuity (e.g. gaps or overlaps) of variation in shape and boundaries across blocks. In local non-affine transformation gaps or overlaps do not occur. The local affine transformation can be constrained so that the variations in shape and boundaries are continuous and smooth between blocks. Thus every set of local transformation (particularly if made continuous across blocks) can be represented by a global transformation, though this representation could be very complex. The local or global divide is therefore an issue of implementation.

A rigid-body transformation requires six parameters (three rotations and three translations). Many papers have reported the use of rigid body transformation for intrasubject intramodality and intermodality registration, making use of measures based on features or the joint intensity histogram. Although rigid-body transformation is predominantly used in the intrasubject registration of brain images, it is also used to align images of other parts of the body where deviations from the rigid body assumption are small, especially at low resolution. Before applying rigid body transformations, especially in multimodality applications, the image field of view, slice orientation, scan geometry and the voxel dimensions should be made consistent between scans.

2.3.2 Affine transformation

The rigid-body transformation is a special case of the affine transformation. In addition to rotations and translations, the affine transformation includes independent scaling of each axis and skew in any direction. Shapes may not be preserved but parallel lines remain parallel. Scaling parameters were used to register retinal images [185]. Other affine parameters, have been used in several applications [186,187,188,189,190]. It is possible to obtain a reasonable registration of most images using just 9 parameters (3 translation distances: t_x, t_y and t_z, 3 rotation angles: φ_x, φ_y and φ_z, and 3 scale factors: s_x, s_y and s_z). The skew is rarely applicable in the case of medical images registration and the three skew parameters are omitted in this discussion. The 9-parameter affine transformation can be represented, in homogeneous coordinates, as a product of 4×4 matrices representing translation **T**, rotation **R** and scaling **S**. The translation and scaling matrices are expressed as

$$\mathbf{T} = \begin{vmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{vmatrix} \text{ and } \quad \mathbf{S} = \begin{vmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

The rotation matrix $\mathbf{R} = \mathbf{R}_x \mathbf{R}_y \mathbf{R}_z$ is a product of matrices representing independent rotations about x, y and z axes, by respective angles φ_x, φ_y and φ_z, where

$$\mathbf{R}_x = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\varphi_x & -\sin\varphi_x & 0 \\ 0 & \sin\varphi_x & \cos\varphi_x & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \quad \mathbf{R}_y = \begin{vmatrix} \cos\varphi_y & 0 & -\sin\varphi_y & 0 \\ 0 & 1 & 0 & 0 \\ \sin\varphi_y & 0 & \cos\varphi_y & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

and

$$\mathrm{R}_z = \begin{vmatrix} \cos\varphi_z & -\sin\varphi_z & 0 & 0 \\ \sin\varphi_z & \cos\varphi_z & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

Using homogeneous coordinates, let $x = (x_1, x_2, x_3, 1)^T$ denote the position of a voxel in the floating image and $y = (y_1, y_2, y_3, 1)^T$ its position following the affine transformation. Although the individual matrices could be combined in different permutations, the following order of operations (right to left) is assumed in this thesis unless otherwise indicated,

$$(y - C_r) = \frac{v_f}{v_r} \times \mathbf{TRS}(x - C_f), \qquad (2.5)$$

where, C represents the image center, v is the voxel dimension and the subscript refers to the floating (f) or the reference (r) image. In global registration, all 9 parameters are varied independently until the similarity measure Φ is at optimum. The overall expression for **TRS** can further be simplified to $\mathbf{Q} = \{q_{ij}\}$ as a function of the affine transformation parameters. **Q** reduces to a rigid-body transformation by setting $s_x = s_y = s_z = 1$.

2.3.3 Non-affine transformation

If an affine transformation is unable to account for variability between images, [186,191,192,193], a non-affine global transformation may be employed (e.g. polynomial transformation [194,195,196,197]. Most non-affine transformations adopt a physical model of tissue elasticity and are expressed in terms of a local displacement vector field. In the physical model, the floating image is modeled as a physical object, such as a thin metal plate [198], a linear elastic solid [199,200], or a viscous fluid [201,202], deformed using forces derived from a local gradient of a potential function. The potential function is computed at a number of control points, such as the locations of anatomical landmarks. In these approaches, local constrained transformation of one image with respect to the second image is allowed. The optimum image transformation expresses equilibrium between external driving forces and internal restoring forces arising within the transformed image. Meyer *et al.* [186] and Slomka *et al*. [203] demonstrated the usefulness of the thin-plate spline model to register thoracic PET and CT images. Modifications suggested by Little *et al.* [204] incorporate rigid structures in the transformation. Rohr *et al.* [205] proposed to relax the transformation in order to account for (anatomical) landmark localization errors. In a further modification to landmark based thin-plate spline, the transformation was made reversible and intensity matching was applied between the floating image and the reference image [206].

Smoothing and regularization techniques have been used in an attempt to overcome the ill-conditioned problems associated with image transformation. Basically, these methods attempt to force smoothness on the solution of a least-squares error problem. In the physical models, regularization in the form of internal constraints induced by the

elastic continuum, is applied to smooth out the local transformation [193,207]. In general, the physical model based transformation is formulated by partial differential equations. These equations are solved analytically or numerically. Christensen *et al.* [201] estimated the transformation based on the finite element approach. Bookstein [198] and Malcolm [200] calculated the solution analytically using basis spline functions. Bro-Nielsen and Gramkow [202] adopted a convolution filter to solve the fluid transformation. Any spatial transformation can be expressed as a displacement vector field comprising displacements of all points in the image. Complex non-affine transformations can be estimated by measuring the displacement of each image voxel with respect to the second image. Figure 2.5 shows a complex global transformation (e.g. rotation) can be approximated by a set of simple local transformations. Estimating the deformation vector at all voxels is computationally expensive and may result in noisy estimates. Dividing the image space into subimages and computing the displacement vector for each subimage can accelerate the registration process. Both the floating image and the reference image are divided into subimages with a uniform 3D discrete grid pattern whose nodes correspond to subimages centers. For each subimage, a displacement vector is obtained that maximizes a local similarity measure. The displacement field is usually filtered to enhance the smoothness of the local transformation [208]. The transformation away from nodes is propagated by spline [193,207,209], radial basis functions [210] or Gaussian interpolation [182,211].

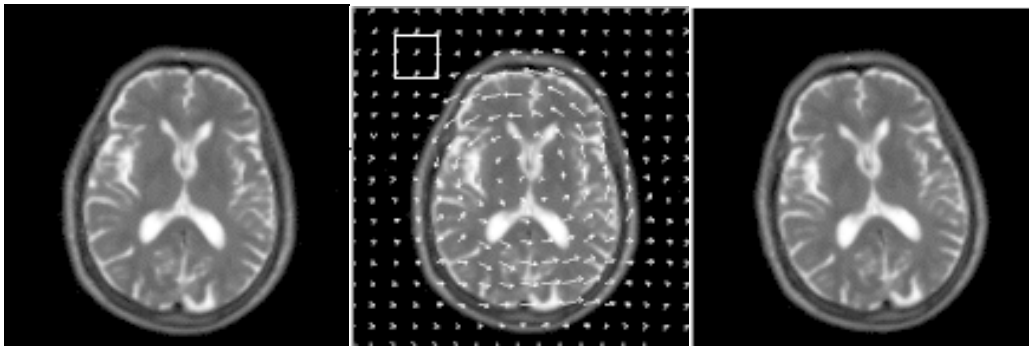


Figure 2.5: A complex global transformation can be approximated by a set of simple local transformations. The floating image (left) is divided into subimages (one is shown as a white box) and each subimage is translated to best match the reference image (right). These local translations can be visualised as the displacement vector field (middle). The field defines the transformation that must be applied to the floating image to align the reference image (illustrated here for a simple rotation).

2. 4 Interpolation of transformed data

The basic problem in registration of an image A (floating) and a second image B (reference) is to find a spatial transformation T such that $T_\theta A$ is as similar as possible to

B. In general, T does not map voxel centres in A to voxel centres of the original grid and interpolation is needed to estimate the intensity at each voxel. Interpolation is often required following a transformation as shown in the 2D example in Figure 2.6. In volumetric imaging, interpolation is often used to compensate for anisotropic data sampling and different voxel sizes. A 3D interpolation is employed to produce isotropic voxels for visual inspection. However, it is undesirable as it increases complexity and overhead without adding new information for image registration.

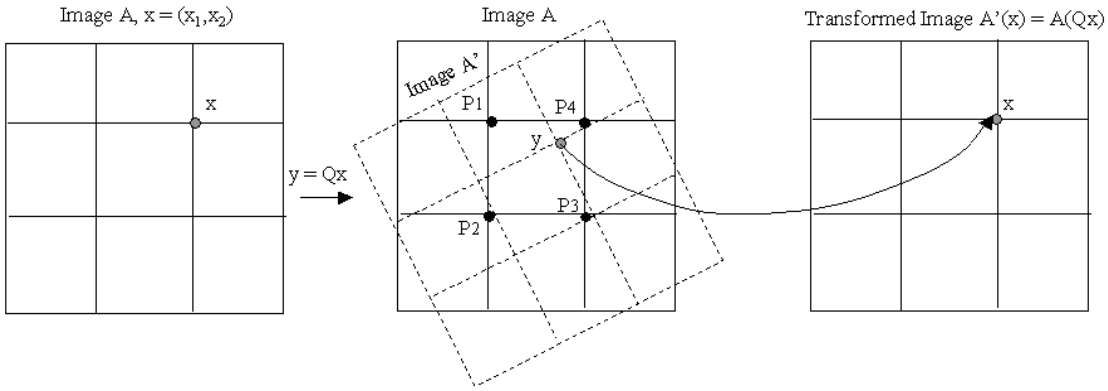


Figure 2.6: Interpolation is often required following a transformation. A 2D transformation T maps every point x in image A to a point y in the transformed version of image A′. To obtain A′, each y is assigned a grey value derived by bilinear interpolation of the nearest 4 pixels of grey value P1, P2, P3 and P4 from A. It shows also the mapping of a point (e.g. bottom left node of the grid) outside the image domain.

Interpolation algorithms are reviewed in several papers [212,213,214,215]. According to Parker *et al.* [216], sampling the interpolated image is equivalent to convolving the image with a sampled interpolation kernel function. The complexity of implementation for image registration depends on the type of kernel function (polynomial, B-spline, Lagrange or Gaussian) and kernel size κ (i.e. the extent of the kernel function in the spatial domain). Let us express an interpolated value A(t) at some continuous non-integer coordinates $t \in \mathrm{R}^3$ from its discrete samples A(x) evaluated at integer coordinates $x \in \mathrm{Z}^3$. The interpolation can be described as the convolution of the κ discrete image samples with the interpolation kernel function, $\mathrm{A}(t) = \sum_{i=1}^{\kappa} \mathrm{A}(x_i) w(t - x_i)$.

Although the sinc function provides the exact reconstruction of A(t), it requires an infinitely large kernel width. In 3D image registration, the most popular approximation to the sinc filter is trilinear interpolation. In nearest neighbour (zero order) interpolation, the intensity value at $t = \mathbf{Q}x$ is assigned the intensity of the spatially closest neighbour.

No matter how many dimensions are involved, there is only one discrete sample that contributes to $A(\mathbf{Q}x)$. It is seldom used for image interpolation due to the severe loss of quality arising from aliasing and blurring. The trilinear interpolation calculates the intensity values at $A(\mathbf{Q}x)$ from the values of eight nearest voxels (i.e. $\kappa = 8$). The trilinear interpolation can be carried out as a sequence of bilinear interpolations in the three orthogonal directions as illustrated in the Figure 2.7.

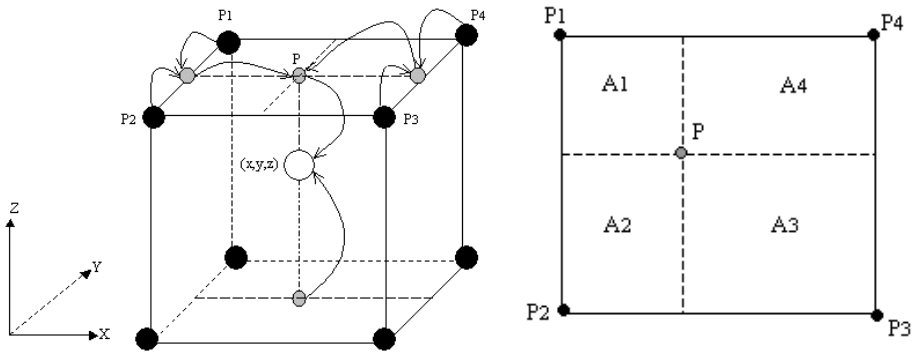


Figure 2.7: (Left) Trilinear interpolation of a point (x,y,z) in a 2×2×2 neighborhood. Interpolation is performed in the y direction first. The intermediate grey values are generated by four 1-D interpolations. This is followed by two 1-D interpolation in the x direction. The resulting two points are colinear with (x,y,z). The final 1-D interpolation is carried out in the z direction. (right) The Figure illustrates a plane view of the top of the cube (left). The main point of (right) is to show weights derived from areas (in 2D) or volumes (in 3D). But since this is a sequence of linear interpolations, the weights are linear at each step. For each interpolation, weights the values at direct neighbours. The grey value at point P is expressed as '$P_1A_3 + P_2A_4 + P_3A_1 + P_4A_2$' where A_i is the partial area in 2D of each of the closest 4 pixels. In a 3D tri-linear interpolation, partial volume is used instead of the partial area for interpolation.

Trilinear interpolation, although not as fast as the nearest neighbour interpolation, is faster than other methods – an important consideration as interpolation is required many times during the iterative registration process. Higher order interpolation kernels, such as the Gaussian [217], cubic [218] and B-spline [213] approximates the ideal sinc interpolator more accurately. However, their larger kernel size increases the computational cost. The kernel of the trilinear interpolation $w(x)$ is stated as

$$w(x) = \begin{cases} 1 - |x|, & |x| \le 1 \\ 0, & \text{otherwise.} \end{cases}$$

A problem with image interpolation is that it blurs the resulting image. This is especially notable if smoothing accumulates through successive transformations. Interpolation can also cause artefacts in the joint-intensity histogram based similarity measures and may prevent image registration from achieving sub-voxel accuracy [219]. Figure 2.8 illustrates this effect as a function of translation for a MR image and a SPET image.

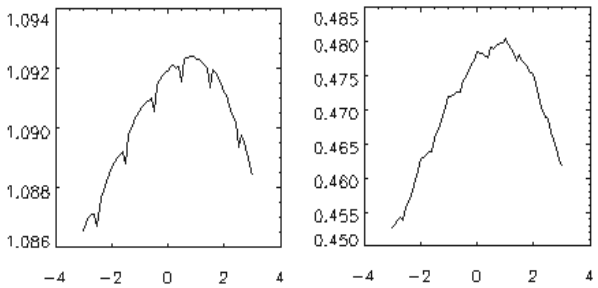


Figure 2.8: Variation of NMI (left) and MI (right) for a MR image and a SPET image as a function of translation in pixels. Full sampling and trilinear interpolation is used for non-integer translations. The true alignment occurs at zero translation. Both similarity measures manifest local maxima. The vertical axis refers to the value of similarity measures of NMI and MI respectively.

Because interpolation alters intensity values, even a small variation in the value of transformation parameters may change the histogram of the floating image leading to unpredictable changes in the dispersion of the joint intensity histogram. Trilinear partial volume (PV) interpolation has been suggested to reduce this problem [220,221]. PV is a normal trilinear interpolation in image space - only the way of expressing the resulting value as a vector of coefficients is different. It is a modified method of estimating the voxel count of a cell in the joint histogram. PV updates the joint histogram fractionally by adding the corresponding weights w_i (not a count of one) corresponding to a transformed location $\mathbf{Q}x$ and its eight neighbours corresponding to the histogram cells $\left(\mathrm{A}'(x_\kappa), \mathrm{B}(x)\right)$. The weights are the trilinear interpolation weights associated with the eight nearest neighbours around $\mathbf{Q}x$. The shape-based interpolation [222], which takes into account the shape of the image intensity histogram, interpolates more accurately than intensity interpolation and warrants further investigation. Where a transformation maps points outside the target image domain, interpolation must give way to extrapolation. The large errors associated with extrapolation make it unattractive for medical image registration. Instead, computation of the similarity measure is confined to the volume of overlap between the transformed floating image and the reference image. It is therefore important to detect points transformed outside the overlap volume.

2. 5 Optimization

Most registration algorithms require an iterative approach to search for the optimal transformation. One exception is the point-based rigid-body registration, where the optimal transformation to align corresponding points of the floating image and the reference image can be computed in one step by singular-value decomposition

[48,223,224]. The iterative optimization routine searches in a parameter space whose dimension is defined by the required number of parameters. A rigid body transformation has six degrees of freedom (corresponding to 3 rotations and 3 translations) giving a six-dimensional parameter space. A full affine transformation requires a twelve-dimensional parameter space. Every time transformation parameters change, a new iteration is performed until an optimal similarity measure Φ is reached. The iterative process continues until it satisfies a stopping rule, which may be expressed in terms of the rate of change of the similarity measure or the parameter values, or a preset number of iterations. Press *et al.* [127] describe optimization algorithms in general and Maes *et al.* [225] compare the performance of various optimization methods in the context of image registration using mutual information.

Optimization methods are classified into two main categories based on whether they require the function's derivatives. Of the methods used in image registration, Powell's algorithm (used in [145,179,208,220,225,226]) and the simplex method (implemented in [182,186,227]) do not require derivatives whereas gradient descent (employed in [177,189,207]) and the Levenberg-Marquardt method do [218]. It is worth noting that these methods are sensitive to local maxima in the parameter space and need to be initialized close to the optimal (unknown) solution. In an ideal situation, the similarity measure is well-behaved as a convex function in the parameter space so that there is a wide capture range for the optimization algorithm to converge to the global optimum. In reality, this is not always the case. The variation of similarity measures in parameter space is generally far from quadratic, and multiple local optima often exist. Noise, coarse resolution and interpolation artefacts make image registration more difficult by generating undesirable local maxima or reducing the value of the global maximum. An optimization algorithm may fail to converge to the global maximum if it becomes trapped at a local maximum. In clinical practice, all image registration should be visually checked to validate the correctness of the registration. Commonly used methods to reduce local maxima include smoothing of the images before registration, rebinning the joint-intensity histogram, and using a multi-resolution strategy.

An alternative approach to circumvent the problem of local optima is by stochastic methods, such as the simulated annealing and the genetic algorithm, which theoretically guarantees convergence to a global optimum. In simulated annealing (used in

[185,228,229]), optimization is taken as a cooling process which is deliberately slow to ensure a global optimum. In the genetic method (employed in [184,229,230,231]), optimization is modeled as the evolutionary survival of the fittest and needs to evaluate a large number of fitness functions. These two methods are computationally expensive and converge slowly. To speed up convergence, Jenkinson and Smith [192] proposed a pseudo global optimization method and applied this to intermodality registration with an intensity-based similarity measure. A coarse grid is initially found over three rotation angles and at each grid point optimal translation is found by local optimization. Registration is refined on a smaller grid by evaluating the lowest cost function over all grid points.

2. 6 Validation

Developers and users of registration methods have measured the accuracy and precision of their methods in many ways. The validation is, in general, a complex issue that is nearly impossible to characterize. Differences in scan protocol, sample size, anatomic site, implementation strategy, intended application, and clinical requirements have made it difficult to compare the registration algorithms quantitatively. Thus, it is necessary to provide a basis for the selection of a registration algorithm for use in a given application. The validation of an image registration algorithm focuses on three main elements: evaluation criteria, metrics and the data sets for evaluation. They are interrelated and should not be discussed in isolation.

2.6.1 Criteria for evaluation

A good image registration technique should be able to demonstrate its "goodness" in terms of accuracy, precision, robustness and reliability. Accuracy refers to how close a measurement comes to its ground truth. To define the true registration is non-trivial. In the clinical uses of registration, ground truths are unknown and it cannot assess the absolute accuracy of different implementation strategies. In the absence of ground truth, the overall registration accuracy, or at least accuracy at a specific image location, could be assessed quantitatively by comparison to some *a priori* superior reference (the gold standard). West *et al.* [122] used a fiducial-marker based method as the "gold standard" to evaluate registration accuracy of a group of retrospective registration techniques applied to clinical data. By quantifying registration accuracy, it is possible for a clinician to differentiate among registration techniques for a particular clinical

application. However, such a gold standard is only applicable to rigid body registration. In validating non-affine body registration, one needs to define true values locally but the issue is still controversial.

Precision indicates the variation of registration errors as the registration algorithm brings the image close to the correct alignment but fine residual error remains. In this case, the error can be measured as the distance between corresponding control points (e.g. fiducial markers, anatomical landmarks, vertices of a bounding box, or randomly generated voxels) in the two images after registration. If the standard deviation of the registration error is small, the registration is regarded as precise.

Robustness refers to the ability of a registration technique to give consistent results regardless of the initial starting conditions. A robust registration technique is expected to converge to approximately the same result with a high success rate on all trials regardless of the variation in image conditions [192,232,233]. A good registration technique also needs reliability to provide a good performance on registration, independent of the image data set and scans protocols so that the results are easily transferable to different clinical applications. To make meaningful comparisons of results from different studies, all criteria for accuracy, precision, robustness and reliability should be reported. However, these criteria are difficult to compare among different reports without referring to the metrics and the image data set used for registration.

2.6.2 Metrics for evaluation of registration methods

A priori reference

The Euclidean distance with respect to *a priori* gold standard is the most common choice of metric for evaluation of registration techniques. It measures the length of the spatial displacement between corresponding voxels in two images after registration. The measurements can be expressed in terms of the mean, maximum, median or root-mean-square (RMS) displacements over the voxels in a region of interest. The advantages of the spatial distance for the comparison are that (1) it gives a direct indication of the registration accuracy in mm; (2) is relatively simple for investigators to provide an accurate identification of the markers; (3) is adaptive to inter or intra-modalities; (4) can

be applied to a group of images. Fiducial markers, anatomical landmarks, vertices of a bounding box and randomly generated voxels within a volume of interest have been reported for use in validation. Fiducial markers, visible in both images, are attached to the subject's skin or mounted to a stereotactic frame. The RMS distance between the corresponding markers taken over all markers gives the fiducial registration error (FRE). However, the measurement of FRE is subject to the so-called fiducial localization error (FLE) caused by marker movement between scans and uncertainty in localization of the markers. The main drawback of this approach is that it needs special preparation with addition of identifiable markers during imaging. It is not suitable for retrospective registration. For most clinical applications, the most important error measure is the target registration error (TRE), which is the RMS distance after registration between corresponding points of interest (target) other than the centroids of fiducials [122,123,234,235]. In most cases, internal anatomical landmarks are used as the target to calculate TRE. The characteristics of TRE in point-based rigid-body registration have been reported extensively [125,236,237,238]. FLE and the markers' configuration were considered to be important factors governing TRE [236]. Grachev *et al.* [128] used anatomic landmarks to assess the accuracy of two popular non-affine intersubject intramodality registrations [239,240,241]. Figure 2.9 shows the difference between the FRE and TRE to measure registration performance. The fiducial markers in general appear to be skin markers and the anatomical landmarks are usually deep interior inside body. This is often the case but anatomical landmarks also appear on the body boundary (e.g. bony protrusions). Perhaps the more important difference in the FRE and TRE is its extrinsic and intrinsic nature of the registration accuracy. Intrinsic errors use body tissues so they give a true measurement of performance whereas extrinsic errors may not. However, these two errors are related as reported in [236].

Another approach to validating a registration technique is based on the RMS distance measured over all vertices of a bounding box compared to a gold standard (e.g. fudicial markers) after registration [133,221]. The maximum displacement error of the vertices provides an upper bound on the registration error within the volume of the box. The registration accuracy can also be calculated in terms of spatial displacement between a set of corresponding voxels in the floating image and the reference image selected randomly over the region of interest before and after registration [233,242]. An alternative approach is to define the accuracy of registration in terms of the difference

between the registration parameters and the gold standard registration parameters. The registration error is expressed in mm for translational errors and in degrees for rotational error [232,243,244]. It does not give the registration error directly in term of the spatial distance. There is also the issue of noncommutativity (changing the order of rotations may lead to different outcomes) and nonuniqueness (the same orientation could be obtained with different parameter values). Small values indicate that the registration is close to the gold standard estimate.

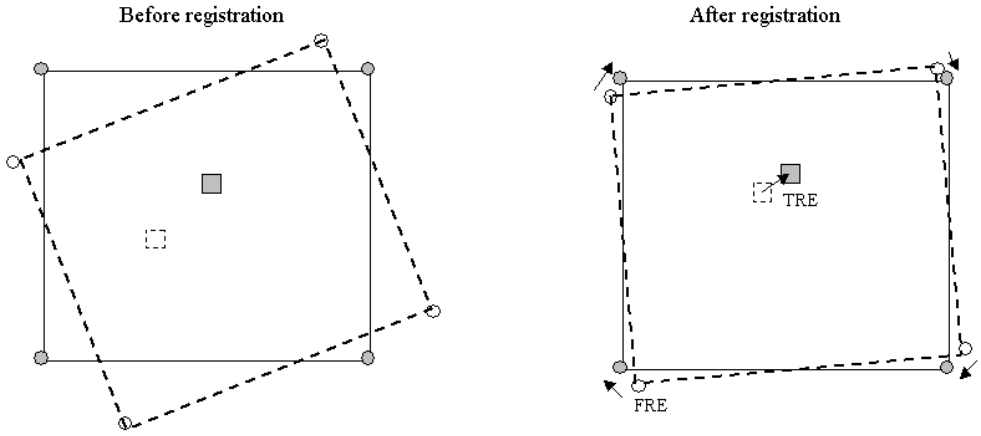


Figure 2.9: Illustration of FRE (extrinsic) and TRE (intrinsic) registration errors. Perhaps the more important difference is in the extrinsic/intrinsic nature of the markers used to estimate the errors. Intrinsic markers are body tissues so they give a true transformation whereas extrinsic markers may not. FRE is measured at each fiducial as the distance between the corresponding fiducial markers (circles) in the reference image and the floating image after registration. TRE is measured as the distance between the corresponding anatomical landmarks (squares) in the reference image and the floating image after registration.

Consistent transformation

As happens to real data sets, fiducial markers are often not available for the gold standard validation. Woods *et al.* [239] suggested the use of internal inconsistency measures as an alternative approach to validate registration. Given three scans (A,B,C) of the same subject, registrations between image pairs are performed among the three images with transformations T_{AB}, T_{BC}, T_{CA} corresponding to the registration of A to B, B to C and C to A. In an ideal situation of perfect registrations between image pairs, the combined transformation $T_{CA}\, o T_{BC}\, o T_{AB}$ maps image A back to its original orientation. The composition of these transformations is an identity transformation **I**. Any deviation from **I** indicates a discrepancy resulting from misregistration (known as error transformation). Internal inconsistency (in mm) is measured by multiplying the error transformation with the position vectors of some selected voxels of the images to compute global mean of the summed squared displacement error. It can be argued that

internal inconsistency is a measure of registration accuracy. Since it does not require fiducial markers, some authors suggest its use for evaluation of registration accuracy [239]. Holden *et al.* [242] used it to compute the registration accuracy over serial MR images. An alternative approach is to multiply the recovered transformation matrix by the inverse of the known transformation matrix at different known positions to compute the error transformation [57]. A practical issue of this approach is a need for 3 images of a single subject. This technique does not give the registration error directly in terms of the target registration error. Further study in this area is warranted to assess target registration error relative to the consistency approach and the fiducial marker based method.

Visual assessment

Accuracy can also be measured qualitatively by visual assessment. For the qualitative assessment of registration accuracy, visualization tools such as dual cursor display [245] or fused images [146,245,246,247] are used to provide the clinician with sufficient information to verify the goodness of registration. This approach relies on the clinician to reject or accept a registration outcome based on a subjective judgment. Clinically, visual inspection insures against severe misregistrations that most algorithms produce in a small number of studies for subtle reasons. Unfortunately, visual assessment is an extremely labour-intensive process. It suffers from both intra and inter-observer variability, which make the definition of correctness in registration difficult. Despite such variability, the sensitivity and reliability of visual assessment is the subject of study by various research groups. Ardekani *et al.* [131] deliberately misregistered MR and PET brain image pairs by translational and rotational displacements about all three axes separately relative to a gold standard. The minimum displacement corresponding to "definite misregistration" as perceived by radiologists, was recorded. The study showed that a trained clinician could detect PET-MR misregistrations of 2.5° rotation about the x- and y-axes, 2.0° rotation about the z-axis, 2 mm in x- and y-translations, and 3 mm in the z-translation. The gold standard for this study was an automatic minimum variance algorithm described earlier [131]. Wong *et al.* [248] also evaluated the ability of a trained clinician to detect PET-MR misregistration visually. With the mutual information algorithm used as the gold standard registration, they showed that a rotational error of greater than 4.0° and a translational error of greater than 3 mm were

reliably detected by a trained observer. A more recent study of MR-CT registration shows that a human observer can visually detect errors of 2 mm or greater relative to the fiducial algorithm as the gold standard [223,224]. These studies suggest that registration error of 2 mm is the limit of an observer's ability to detect a misalignment in intrasubject intermodality brain registration for most clinical data sets. It is interesting to note that the Figure is comparable to the registration accuracy of voxel-based algorithms compared to fiducials registration in [123]. Median target registration errors in the range of 2.0-3.6 mm were reported by West *et al.* [122] for MR-PET registration using Wood's algorithm and mutual information.

2. 7 Image data set

Validation of *in vivo* studies is complicated by the lack of ground truth. Apart from clinical images, synthetic images and physical phantoms were used in a large number of studies for validation. Ordinary MR or CT images of a real subject correlated with post mortem sections can be used to synthesize a SPET dataset (e.g. Zubal phantom [249,250], Visible Human data set [251], digital brain phantom [252]). The synthetic image data set is controlled in almost every aspect to simulate images acquired in a clinical situation. Furthermore, registration with simulated data permits controlled evaluation over a wide range of conditions such as contrast, spatial resolution, scan geometry, field of view, slice orientation, noise, scatter and non-uniform attenuation. Using digital phantoms to assess the performance of PET-PET [253] and CT-SPET [254] registrations have been developed and employed. In simulation, ground truth is known and accurate gold standards can be produced. To validate MR-PET registration techniques, one approach is to simulate PET images by segmentation of MR images followed by the assignment of different uptake values to various tissue types. The simulated PET images are then registered to the MR images using various initial starting transformations so that the exact (rigid body) transformation between each MR image and its associated simulated PET images is known. The registration results are compared to the known transformation in order to evaluate the registration accuracy [253,255,256,257,258]. Since the same anatomical phantom is used to mimic different modalities, it can be used to test intermodality registration algorithms. However, the digital phantom fails to model some factors that contribute to scan-to-scan variability in clinical data, such as non-rigid deformations due to respiratory or cardiac motion, distortions due to non-rigid movement of internal organs during scan acquisition and

artefacts due to field inhomogeneities in MR. Image data for validation can also be generated by a physical phantom. The merit of the physical phantom is that it is subject to true image acquisition with all experimental limitations and artefacts. Moreover, activity uptake in the various "tissue" compartments can be independently controlled. These phantoms include the Hoffman brain phantom [259] simulate uptake in the brain, the anthropomorphic cardiac phantom [260] with lung, heart and spine inserts for cardiac SPET imaging and the torso phantom [186,261] to simulate normal patient's abdominal-thoracic imaging. Cold and hot "lesions" can be simulated in the scans. To obtain the ground truth, markers are often attached to the phantom for use as fiducials in quantitative assessment of registration accuracy. The downside of the simple physical phantoms is that they do not adequately model the anatomical and physiological complexity of the human body. Furthermore, the design of any physical phantom requires a long-term chemical stability and reproducibility, which is difficult to achieve. Clinical data are therefore needed to evaluate registration accuracy. One critical issue that will be important for the future acceptance of these approaches to validation will be the choice of the data set. The data should include a wide range of image quality, a large number of subjects, as well as a variety of modalities and acquisition regimes within each modality. The data should be made available to all investigators. Since most studies depend implicitly on data sets to test for accuracy, reliability and robustness, this makes it easier to repeat implementation and to replicate image acquisition parameters, but on different subjects, for objective evaluation.

Chapter 3

Joint-intensity histogram techniques

The purpose of this chapter is to present joint-intensity histogram based image registration techniques. In these methods, registration is based solely on the intrinsic properties of intensity or grey-scale values of volume elements of the two images to estimate the degree of similarity between them. These recent developed techniques are derived from communication and statistical theories endeavouring to use the full content of the image information involved. One distinct capability of such matching criteria is robustness and registration performance for multimodality image registration.

The concept of image intensity as random variables in the context of registration is introduced. The probability density distribution and the joint probability density distribution of these random variables are estimated from a 2 dimensional plot showing the combinations of intensity values in all the corresponding points of the two images. In the case of information theory based methods, image registration is viewed as a communication in a noisy channel with an input signal (floating image) and an output signal (reference image). An optimal registration is obtained when the relative redundancy of information of the input and the output signal is minimized. Signal redundancy is measured by joint entropy. Thus the entropy as well as the joint entropy of images and their measurement is introduced. The other approach is to use simple statistical measures (e.g. mean, variance and conditional variance) for image registration. These methods make use of a general observation that, in spite of the different intensity sets that represents various tissues, uniformity of physical properties (such as attenuation coefficient in CT, proton density in MR and tracer concentration in SPET or PET) measured in image partitions is a robust similarity measure. The layout of this chapter is to introduce the notation that an image voxel can be taken as a random variable for the purpose of image registration. It follows with the description of the joint intensity histogram to estimate probability distribution and joint probability distribution of intensity gray values and a discussion of the entropy-based and variance-based image registration. Adaptation and modification of these algorithms are also presented.

3.1 Images as random variables

Despite the different modalities for image acquisition, typical medical images are reconstructed in the form of a 3D matrix or grid consisting of discrete volume elements. Each element creates a 3D volume element, or "voxel", which is shaped like a little box. This is depicted in Figure 3.1. In the case of isotropic mode of image acquisition, equal dimension voxels are obtained in all three orthogonal planes. In the anisotropic mode, one of the voxel dimensions is unequal to the others. For example, the spacing between contiguous axial planes along the z direction may be larger than the voxel size in the transverse x-y plane. In this case, length along the axial direction is larger. The intensity value(s) that represents a voxel depends on the imaging modality, signal intensity and the tissue type involved from the corresponding location. The intensity value of each voxel gives an average signal acquired from tissue within the corresponding voxel. For example, bones may appear with low intensity in MR images but with high intensity values in CT images.

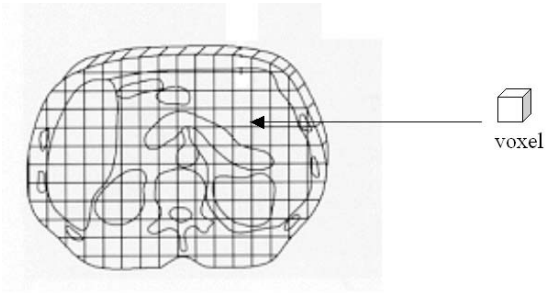


Figure 3.1: Each discrete 3D volume element in the image represents a voxel of tissue.

For the sake of image registration, it is assumed that each voxel expresses a random variable that is quantized to some discrete set of values called intensity or grey-scale. An important consequence of treating images as random variables is that a statistical framework can be used to investigate the image registration problem. To simplify the description of how the intensity of each voxel is reached, each voxel value is treated as a mapping from the 3D matrix volume elements onto a finite set of discrete values $\{0,1,2,\ldots,M\text{-}1\}$ that represent intensity. For an 8-bit encoded image, M equals 256. Thus each voxel may take an intensity (grey) value from this set of number in the range 0 (black) to 255 (white) for 8-bit encoding. All voxel values are assumed to be independent with the same probability density function (PDF). For a 3D image having a sufficiently large number of voxels, the PDF of intensity values can be approximated by histogram measurements. By examining a large collection of voxels, the histogram

estimate of the PDF is the frequency ratio [262]. In an image A the PDF of intensity $a \in \{0,1,2,...,M-1\}$ is denoted by P(a) which is given by

$$\mathrm{P}(a) = \frac{\mathrm{N}(a)}{\mathrm{N}}, \tag{3.1}$$

where N(a) represents the total number of voxels having intensity "a" and N denotes the total number of voxels. Figure 3.2 illustrates an intensity histogram computed from a MR image. The use of PDF makes an idealized assumption that all voxels with a particular intensity represent the same subimage. As in Chapter 2, the term "partition" refers to the set of subsets (rather than an individual subset) in the rest of this chapter. Thus intensity labels each voxel of the image according to which partition it belongs. The image is partitioned into M "disjoints" subsets, the size of each giving the probability of the corresponding intensity being measured. Thus intensity has a probability associated with it.

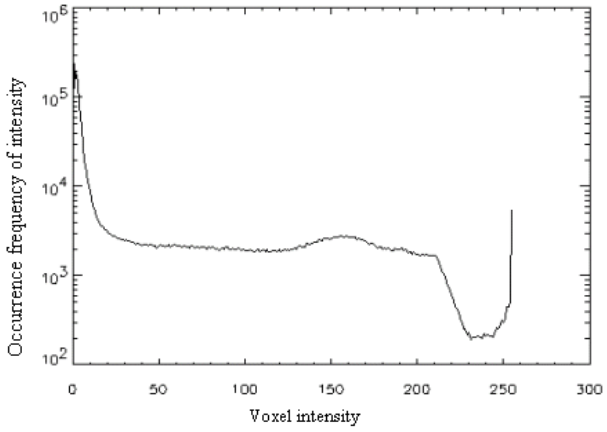


Figure 3.2: Example of an intensity histogram for a MR image is illustrated. The vertical axis denotes the occurrence frequency of intensity and is scaled logarithmically. The horizontal axis refers to the intensity values.

The concept of PDF of a single image can be extended to the joint PDF of two images. Assuming that images A and B are comprised of independent voxels in three dimensions, one may use the same approach to treat the voxel intensity of the images as random variables with associated PDFs. The PDF of intensity in A and B are denoted by P(a) and P(b), respectively, where $a,b \in \{0,1,...,M-1\}$. Here A and B are assumed to have the same discrete grey-scale values. The joint PDF P(a,b), denoting the probability of occurrence of intensity pair (a,b) with intensity $a \in \mathrm{A}$ and $b \in \mathrm{B}$, is approximated by

$$P(a,b) = \frac{N(a,b)}{N}, \qquad (3.2)$$

where N(a,b) is the number of occurrences of (a,b) and N denotes the total number of voxels. This concept is useful for image registration since all voxel-based similarity measures rely on the fact that there is a statistical relation between intensities of the corresponding voxels in the images [48]. In order for this method to work, statistical methods must be used for the computation of the relationship between P(a) and P(b). To estimate the PDF and the joint PDF of images is non-trivial. A Parzen window [189] and joint intensity histogram [179] were used to approximate the joint probability density function. The so-called Parzen estimate of PDF is computationally more expensive than the histogram approach. Thus the Parzen method is not considered in this thesis. The joint intensity histogram approach is easy to implement and gives a visual interpretation of image alignment. It will be discussed in the next section.

3.2 Joint intensity histogram and its interpretation

To visualize the dependence between the transformed image A′ and the reference image B, a 2D joint histogram is created by counting the number of co-occurrences of each intensity pair (a,b) of intensity $a \in A'$ against the corresponding voxel of intensity $b \in B$ for all voxels over the overlapping image volume of A′ and B [173,175,176]. The 2D joint histogram involves a spatial transformation T, which maps A to A′ by $A' = T_\theta A$. Let $A'(x) = T_\theta A(x) = A(\mathbf{Q}x)$ where the transformation matrix $\mathbf{Q}$ maps an element of the domain Ω_A of image A to the domain Ω_B of image B. Consider the value of a voxel x of the reference image to be b and $B(x) = b, x \in \Omega_B$. The location of the voxel x is transformed by $\mathbf{Q}$ to x' (i.e. $\mathbf{Q}x = x'$) and the corresponding value in the transformed image A′(x) is given by

$$a = \begin{cases} A(x'), & x' \in \Omega_A \\ \sum_{y \in n} w_y A(y), & \text{otherwise,} \end{cases}$$

Where $n \subset \Omega_A$ is a neighbourhood of x', $n = \{y \in \Omega_A : \| y - y' \| \le \varepsilon\}$, over which a weighted average is computed with weights w_y. That is, if the transformed voxel position does not lie on the grid of A, its value is obtained by a local interpolation. The intensity is computed over the overlap region of B and A′ of the grid $\Omega_{AB} = \Omega_{A'} \cap \Omega_B = \Omega_{TA} \cap \Omega_B$, and $x \in \Omega_{AB}$. Each point (a,b) in the joint histogram

represents the number of voxels N(a,b) that have an intensity "a" in the transformed floating image A′ and an intensity "b" in the reference image B. A joint PDF of intensity can be estimated from the 2D histogram by dividing each entry N(a,b) in the histogram by the total number of voxels N as in Equation (3.2). The PDF and marginal PDFs of the intensity for image A and B are given by $\mathrm{P}(a) = \sum_b \mathrm{P}(a,b) = \frac{1}{\mathrm{N}} \sum_b \mathrm{N}(a,b)$, and $\mathrm{P}(b) = \sum_a \mathrm{P}(a,b) = \frac{1}{\mathrm{N}} \sum_a \mathrm{N}(a,b)$, respectively, where P($a$,$b$) is the joint PDF. The marginal probabilities P(a) and P(b) are effectively determined from the 2D joint PDF by summing the entry of P(a,b) on the two orthogonal axes rather than by direct computation from the voxels in the images. Figure 3.3 depicts the formation of the joint histogram. In general, the 2D intensity histogram depends on the size of overlapping region Ω_{AB}. It is important to note that the marginal probabilities implicitly depend on both the size and location of overlap between the two images.

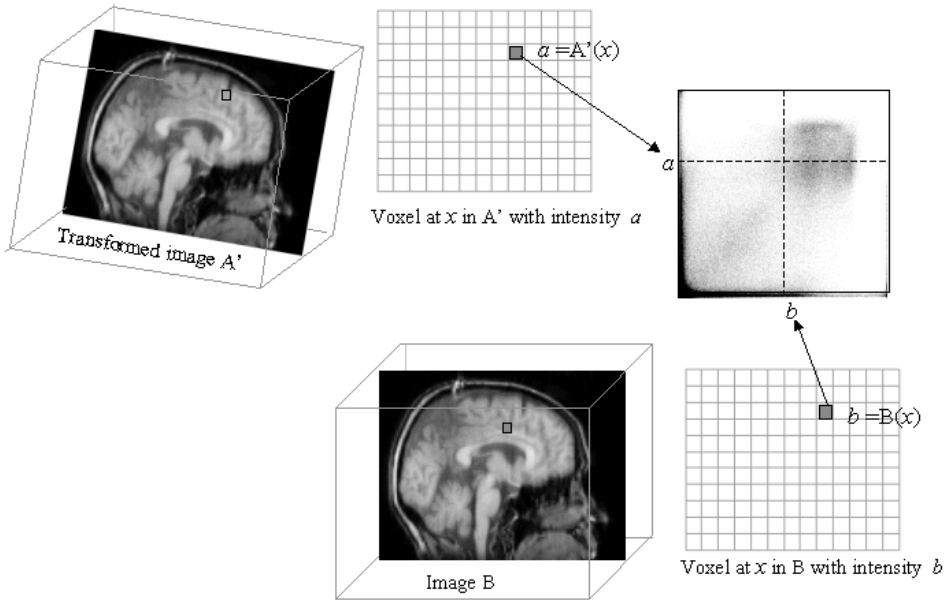


Figure 3.3: Formation of the joint histogram of a MR image with itself is shown. Two images A (floating image) and B (reference image) are defined over discrete 3D grid. The intensity pair $(a,b) = (\mathrm{A}'(x), \mathrm{B}(x))$ refers $a \in \mathrm{A}'$, $b \in \mathrm{B}$ and $x \in \Omega_{AB}$. The intensity a of A′ is related to A via a spatial transformation T. In registration, transformation T maps A to A′ where $\mathrm{A}'(x) = \mathrm{T}_\theta \mathrm{A}(x) = \mathrm{A}(\mathbf{Q}x)$. A′($x$) may not coincide with the grid point of Ω_A and local interpolation is needed to compute the intensity of A′ at x. See text for details.

The centre of voxels of the transformed image T_θA may not coincide with the center of voxels in image A. Intensity interpolation is needed at grid Ω_A in image A to obtain the

intensity values at A′(x). It may also happen that interpolated positions may fall outside the bounds of the image A. That is, elements of the x, y, or z arguments that are either less than zero or greater than the largest subscript in the corresponding dimension of A. The interpolated intensity values can be set either equal to the value of the nearest voxel of A or to a user-specified value. This introduces two practical considerations. First, interpolation involves blurring and new intensity values may be introduced in the resulting image A′. The new intensity values may affect the histogram and joint PDF, producing a negative effect on the performance of registration [48]. Second, T_θA may fall outside the bounds of the image A. The joint histogram is defined only for $x \in \Omega_{AB}$. Interpolation is meaningful if the transformation position x′ (i.e. $x' = \mathbf{Q}x$) is inside the overlapping region Ω_{AB}. Steps must be taken to avoid histogram computation outside Ω_{AB}.

The dependence of A on B is quantified, with the aid of the joint intensity histogram, using a similarity measure. The histogram changes as the transformation parameters change. A general observation in the joint histogram is that intensity pairs seem to disperse when the registration deviates from an optimal registration [48]. However, with an improvement of alignment, the distribution of intensity pairs in the joint histogram tends to show more peaks as well as fewer clusters. When the images are registered, intensity values of the particular anatomical structures, which they both represent, are aligned in the joint histogram. For example, peaks with less dispersion are observed along diagonal direction in the joint histogram as shown in Figure 3.4 when MR and SPET images are registered. When the images are misaligned, the joint histogram shows dispersion or “blurring” because an intensity value in one image is mapped to a large range of intensity values corresponding to different tissue types in the other image. The greater the misalignment, the more dispersion is observed as the intensity cluster becomes more contaminated by the intensities from other anatomical structures. Based on this observation of the characteristic of the joint histogram of two images, quantitative measures of dispersion emerged for use in image registration.

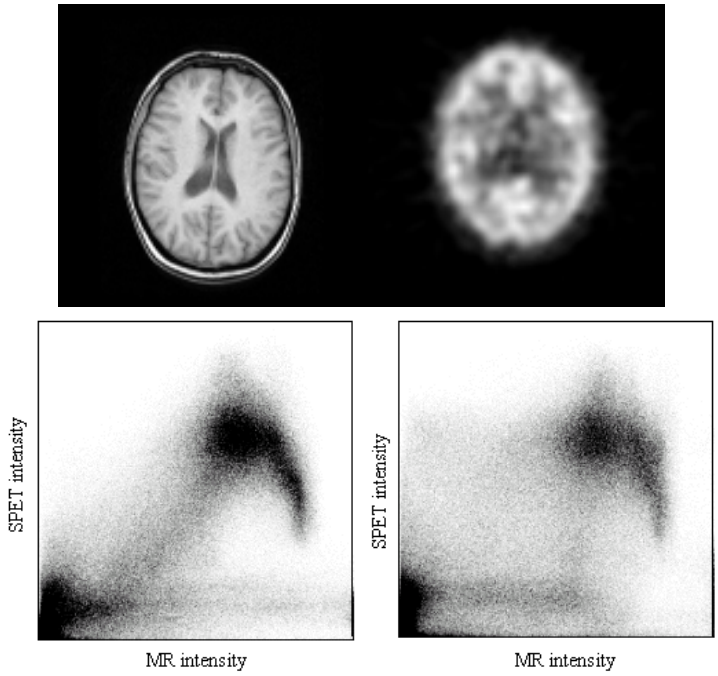


Figure 3.4: A 2D intensity histogram for a MR (top left) and a SPET (top right) is given. The histogram (bottom left) arises when the images are correctly registered. The histogram at the bottom right shows the results when the images are misaligned.

3.3 Measurement of dispersion

All the voxel based similarity measures evaluated, attempt to quantify the observed variation in the 2D joint histogram with spatial transformation to account for the change of dispersion [175,178,179,232]. Certain quantities such as variance and entropy are used to express the statistical relationship between two images. The formulation of each measure makes assumptions about the nature of the relationship. These, in turn, are affected by how the image registration problem is defined. Various similarity measures for use in image registration will now be formulated and their assumptions examined. The most general assumption in image registration is that some form of a *relationship* exists between intensities corresponding to the same anatomical structure in the images A and B [48]. When the images are registered, they should verify that relationship. As pointed out by Roche *et al.* [263], dependence of the images intensity can be modeled in a number of different ways using identity, affine, functional or statistical relationships. The specifics of the registration problem will dictate the form of the relationship and the choice of the similarity measure. Image registration can also be viewed as a maximum likelihood problem [263,264]. Given the intensities in one image as the input signal, the problem of image registration is to reconstruct the intensities in the corresponding voxels of the other image as the output signal with minimum

probability of error. In other words, the optimum transformation for image registration is one that minimizes the likelihood of uncertainty, confusion or variability in the output intensity of the floating image based on the input observable intensity of the reference image. Thus when two images are registered, one image is a good predictor of the other image with minimum uncertainty. If image registration is treated as an intensity mapping between images, then by analogy with communication systems, as depicted in the Figure 3.5, the images are registered when the communication channel between A and B is optimal. The idea of image registration based on information theory was proposed by Collignon *et al.* [176,220] and by Viola and Wells [117,177,189], and was applied with great success in rigid-body multimodality registration problems. This technique will be referred to as the entropy-based image registration approach. Another approach, suggested by Woods *et al.* [172], Ardekani *et al.* [131] and Roche *et al.* [181,265], is based on variance. Image registration is optimized by minimizing the variance of the reconstructed intensity in one image with respect to partitions defined in the other image. Woods' approach (described in Subsection 2.2.2) partitions the MR image based on intensity values, Ardekani's method (also in Subsection 2.2.2) uses segmentation to define spatially connected partitions, and Roche's method (to be described in detail in Subsection 3.5.3) is similar to Woods' but uses a different normalization. It is worth noting that because both the Woods' and Roche's methods sample the histogram, they can be interpreted in terms of joint histogram while Ardekani's method probably cannot.

3.4 Entropy-based algorithm

3.4.1 What is entropy?

Entropy is a measurement of information content of a signal that arose from communication theory in the 1940s [266,267]. Shannon was a mathematician who worked on problems in signal transmission within a communication system as shown in the Figure 3.5. He was concerned with the amount of information carried by a signal rather than with the meaning of the message. He realized that the information contained in the message depends upon the degree to which expectation was confirmed by the receiver and not on the set of symbols that comprise the transmitted message. Suppose that there is a soccer match between two teams, one at the bottom and the other at the top of their division. What is the information conveyed by the two alternative radio

announcements: ‘the bottom team lost’ and ‘the bottom team won’? A listener would not be surprised to hear the first but would be surprised by the second. The first message conveys what he/she would expect to hear (the event was highly likely). The second message would surprise the audience because that event was unlikely to occur.

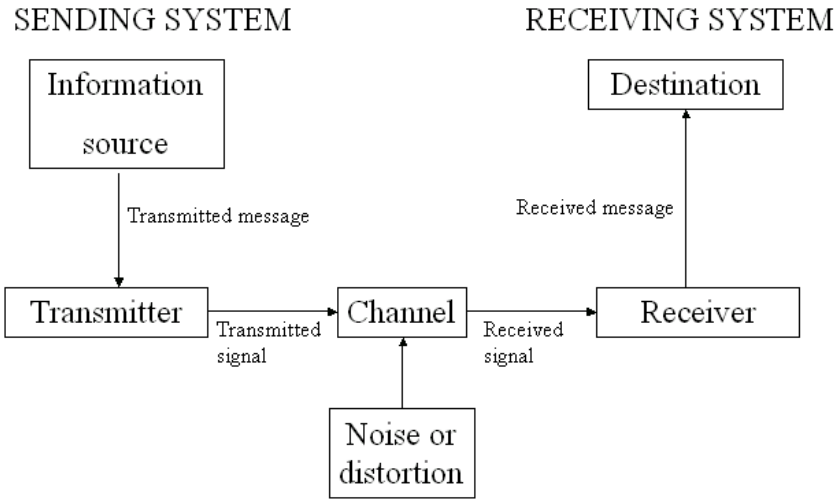


Figure 3.5: A generic communication system

The two alternative messages are equivalent in that they comprise the same number of words, take the same amount of time to say, and so on, yet the information content is clearly different. Shannon [266] proposed to quantify the information about an event having some probability of occurrence by the negative of the logarithm of that probability.

$$Information \propto -\log_2(probability) \tag{3.3}$$

The smaller the probability of an event, the more information is conveyed. If an event is certain to occur, its occurrence conveys no information (-log1 = 0). The more rare the event, the more it can surprise the one who receives the message but, *on average*, the expected information is low because it has a small chance of occurrence. In the real world, it is more meaningful to consider the average amount of information conveyed by the event. Weighting the information per event with its probability of occurrence does this. As a hypothetical example, let us consider a student who performs an experiment to count the occurrence of each number (“1” to “6”) for a fair dice and an unfair dice (i.e. loaded). The outcome is assumed to be a random variable of a set of numbers $\{1,2,3,4,5,6\}$. In the fair dice, all numbers are equally likely to occur. However, for the unfair dice, this is not the case. Let the probability of occurrence of each number in the fair dice be 0.167 (here an approximation of 1/6 has been made) while they

change to 0.95 for "1" and 0.01 for the rest of the numbers from "2" to "6" in the unfair dice. In the fair dice the average information to be gained is equal to 0.298 for each number. However, in the unfair dice the average information is 0.048 and 0.046 corresponding to "1" and "2" to "6" respectively which are much smaller than the amount of information gained from the fair dice for each number. To further illustrate the example, in the unfair dice, the information given by the number "2" is 4.605 (compared to 1.792 in the fair dice) but its amount of information gained on average is only 0.046 (compared to 0.298 in the fair dice). Since not all letters are equally likely to occur in a message comprised of English letters, the probability of occurrence of each letter was taken into account when Shannon defined the information content of a message. He considered a message consisting of *n* permissible letters with probability of occurrence $P_0, P_1, P_2, ..., P_{n-1}$. He weighted the information of each symbol by its probability of occurrence. The Shannon-Wiener entropy measure H is defined as

$$H(P_0, P_1, ..., P_{n-1}) = -\sum_{i=0}^{n-1} P_i \log_2 P_i = -\langle \log_2 P \rangle = -E\{\log_2 P\} \qquad (3.4)$$

Note that entropy H is the mean or expected information of a probability distribution from a certain set of events rather than the individual information associated with a particular event. The resulting entropy of a set of events is the average amount of information to be obtained from the set of their probabilities of occurrence [268]. A logarithmic base 2 is assumed in the rest of this chapter. Entropy is also a measure of uncertainty. Shannon showed that H is at maximum when the probability density function is uniform (all events are equally likely). In this case, it is not possible to tell which event is more likely than any other. In the example of rolling a fair dice and an unfair dice, the entropy of the fair dice is 1.79 while the entropy has dropped to 0.28 when one number is dominant. There is less uncertainty about which number will come up in the unfair dice than the fair dice. When you roll the unfair dice, the number is almost certainly to be "1".

3.4.2 Entropy of an image

As defined in Section 3.1, an image maps a grid of voxels into a set of intensity values $a \in \{0,1,2,...,M-1\}$. The probability density function P(a) of the intensity values in image A can be determined by the number of occurrences of each intensity value in the image and divided by the total numbers of voxels. The entropy of an image is subtle

and it is difficult to explain and describe its meaning. Suppose a set of possible voxel intensity values whose probabilities of occurrence are known. The entropy measures the amount of *information* [266] generated by the image per voxel. It gives the number of bits per voxel, on average, required to represent the image. For example, if an image has a uniform intensity, its entropy is zero. Whenever you pick a voxel, the intensity value is certainly known (no uncertainty or zero information content). In this case, one (2^0) bit is required to store the image. The entropy of an image depends on the number of intensity values (bins). An image with many different intensities has high entropy. When rebinning reduces the number of intensity values, the image has lower entropy. It implies that less number of bits is needed to regenerate the image as the number of bins reduces. For example, when the number of intensity levels in a MR image (Figure 3.6) is reduced from 256 to 2 in steps with a factor of 2, its entropy changes gradually from 3.66 to 0.58.

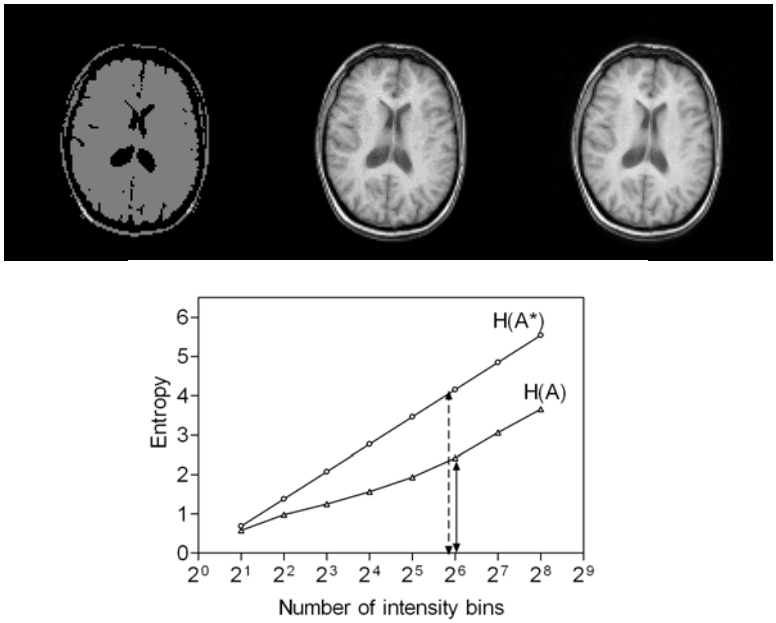


Figure 3.6: Effect of intensity rebinning of an image A (e.g. MR image) on entropy is shown. The images (from left to right) of intensity bins 2^1, 2^4 and 2^7 show the change of grey levels. The plot shows the variation of entropy H(A) with number of intensity bins is by A. The number of bins in the horizontal axis is scaled logarithmically. The plot A* shows the change of entropy for a hypothetical image A* of uniform probability distribution for the intensity value. At 2^6 bins H(A*) referred by the dotted arrow is greater than H(A) denoted by the solid arrow. The smaller entropy of H(A) indicates intensity distribution in A is not truly uniform.

The results show a linear relationship of entropy with the logarithm of the number of intensity bins. The maximum value of entropy is when the probability of occurrence of voxel intensity values is uniform and all probabilities are $1/M$, where M is the number of intensity bins. The hypothetical MR image A* of uniform probability distribution

always generates more entropy than the MR image A (Figure 3.6). It indicates that the intensity distribution of the MR image is not uniform and preference of intensities exists.

Intensity rebinning changes the intensity distribution of an image by repartitioning its intensity histogram. Any change towards a more uniform histogram will increase the entropy [48]. Such changes may result from image smoothing (see Figure 3.7). Paradoxically, noisy images also manifest higher entropy, as noise distributes intensity values about a mean. Smoothing does not necessarily increase entropy. If the image is noisy, smoothing may decrease entropy by suppressing noise. Interpolation also involves smoothing and hence will affect the entropy of the image [219].

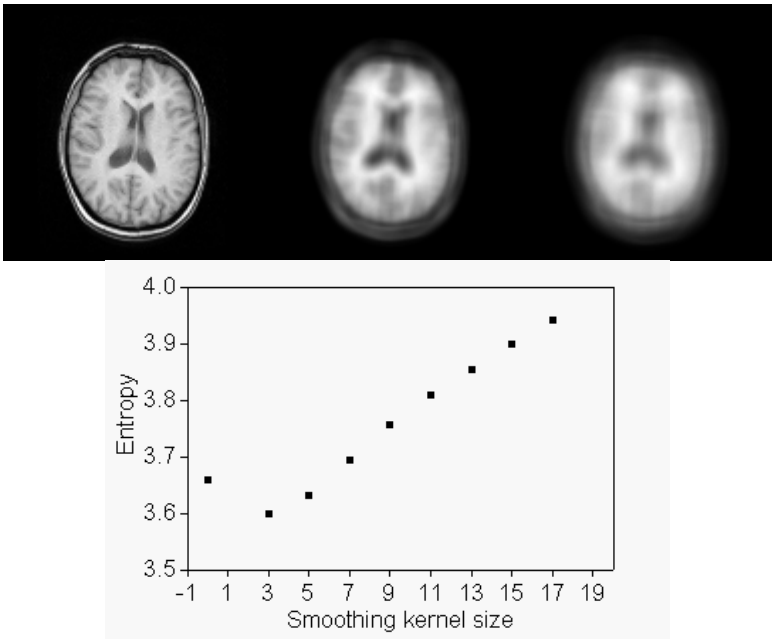


Figure 3.7: Effect of blurring on an image is shown. The left image shows the original MR, and the middle and right images show original MR after using 7^3 and 11^3 boxcar-smoothing filter. The variation of entropy against smoothing kernel size is shown. The entropy dips when the kernel size is 3^3 because smoothing reduces the noise in the image.

3.4.3 Mutual information

The key difficulty in image registration is to determine how one image is related to the other image. One needs to quantify the dependence of one random variable (say, image $\mathrm{A}(x) \in \{1,2,3,...,M-1\}$) on the other random variable (image $\mathrm{B}(x) \in \{1,2,3,...,M-1\}$). Since the information content of an image can be calculated from entropy (uncertainty), then it can be used to quantify the information content of the joint histogram for images A and B. In this section, the joint probability distribution and its relationship to entropy will be considered and how this may be used to interpret the dependence between the

two random variables. In section 3.2, the stochastic relationship between two random variables was described by the 2D joint intensity histogram. After normalization by the total number of histogram entries, the joint probability density function P(a,b) is estimated. The joint entropy H(A,B) is defined as

$$\mathrm{H}(\mathrm{A},\mathrm{B}) = -\sum_{a}\sum_{b} \mathrm{P}(a,b)\log \mathrm{P}(a,b) = -\mathrm{E}_{ab}\{\log \mathrm{P}(a,b)\}. \qquad (3.5)$$

Information in one image is expected to reduce the uncertainty of the other image. This is expected since entropy can be used to relate the uncertainty of two random variables. The joint PDF P(a,b) of two random variables A and B is related to the conditional probability P(a|b) and marginal probability P(b) by the following probability equation

$$\begin{aligned} \mathrm{P}(a,b) &= \mathrm{P}(a \mid b)\,\mathrm{P}(b), \\ \mathrm{P}(b,a) &= \mathrm{P}(b \mid a)\,\mathrm{P}(a), \end{aligned} \qquad (3.6)$$

Where $\sum_{a} \mathrm{P}(a \mid b) = 1$. Since the two joint PDFs P(a,b) and P(b,a) are identical, Bayes' formulation is obtained $\mathrm{P}(a \mid b) = \dfrac{\mathrm{P}(a)\mathrm{P}(b \mid a)}{\mathrm{P}(b)}$. Expanding the expression for the joint entropy (3.5) by (3.6), H(A,B) can be rewritten as

$$\begin{aligned} \mathrm{H}(\mathrm{A},\mathrm{B}) &= -\sum_{a}\sum_{b} \mathrm{P}(a \mid b)\,\mathrm{P}(b)\log\left[\mathrm{P}(a \mid b)\,\mathrm{P}(b)\right] \\ &= -\sum_{a}\sum_{b} \mathrm{P}(a \mid b)\,\mathrm{P}(b)\log\mathrm{P}(a \mid b) - \sum_{a}\sum_{b} \mathrm{P}(a \mid b)\,\mathrm{P}(b)\log\mathrm{P}(b) \\ &= -\sum_{a}\sum_{b} \mathrm{P}(a,b)\log\mathrm{P}(a \mid b) - \sum_{a} \mathrm{P}(a \mid b)\sum_{b} \mathrm{P}(b)\log\mathrm{P}(b) \\ &= -\sum_{a}\sum_{b} \mathrm{P}(a,b)\log\mathrm{P}(a \mid b) - \sum_{b} \mathrm{P}(b)\log\mathrm{P}(b) \\ &= -\mathrm{E}_{ab}[\log\mathrm{P}(a \mid b)] - \mathrm{E}_{b}[\log\mathrm{P}(b)] \end{aligned} \qquad (3.7)$$

For any particular value b that B can assume there is a conditional probability P(a|b) that A has the value a. Shannon proposed to define the conditional entropy $\mathrm{H}(\mathrm{A} \mid \mathrm{B})$ of A given B as the average of the entropy of A for each value of B, weighted according to the probability of getting particular B. The first term in (Eq. 3.7) is the conditional entropy of A given B. This quantity measures on average the uncertainty of A given the knowledge of B. The second term is the entropy H(B) which measures the uncertainty of B. Therefore, the joint entropy of A and B is given by

$$\mathrm{H}(\mathrm{A},\mathrm{B}) = \mathrm{H}(\mathrm{A} \mid \mathrm{B}) + \mathrm{H}(\mathrm{B}). \qquad (3.8a)$$

$$\mathrm{H}(\mathrm{A},\mathrm{B}) = \mathrm{H}(\mathrm{B} \mid \mathrm{A}) + \mathrm{H}(\mathrm{A}). \qquad (3.8b)$$

It can be shown that the equation is symmetrical with respect to the two random variables A and B. The relationships among the joint entropy, conditional entropy and entropy of the two independent variables can best be visualized graphically as shown in Figure 3.8.

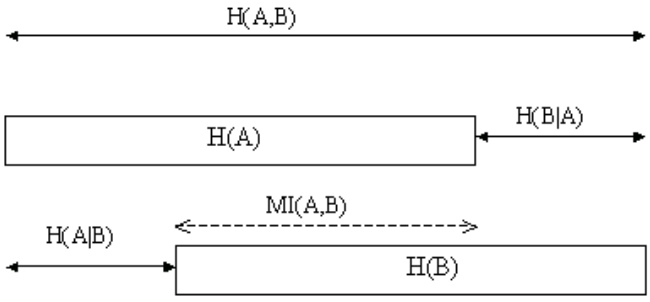


Figure 3.8: A schematic diagram of the relationships involving the joint entropy H(A,B), conditional entropies $H(A \mid B)$ and $H(B \mid A)$, marginal entropies H(A) and H(B), and the mutual information MI(A,B). Note that MI measures the amount of the reduction of uncertainty about one image due to the knowledge of other image.

Equation (3.8a-b) states that the total uncertainty of two random variables (images A and B) is a sum of two terms: the conditional entropy of A given B, which measures the uncertainty of image A when image B is known (and *vice versa*), and the entropy of A, which gives the uncertainty of image A. Note that there is a reduction in the amount of uncertainty (i.e. "choice" of voxel intensity values) about one image given the knowledge of the second image. The amount of reduction of entropy (uncertainty or information content measured by the average number of bits per voxel) is an important measure that quantifies the degree of dependence between the two images. The measure of this reduction in uncertainty of one image given the other image is known as mutual information (MI) and its most frequently used definition is

$$MI(A,B) = H(A) - H(A \mid B), \tag{3.9a}$$
$$MI(A,B) = H(B) - H(B \mid A). \tag{3.9b}$$

It is not difficult to recognize that A and B in Equation (3.9) are interchangeable. From Equations (3.9) via (Eq. 3.8), MI can further be expressed in its second definition

$$MI(A,B) = H(A) + H(B) - H(A,B). \tag{3.10}$$

Note that, if H(A) and H(B) are fixed, maximization of MI(A,B) is equivalent to minimization of H(A,B). Consider image registration as an optimization problem that aims to maximize the amount of reduction of uncertainty of the transformed floating image A′ when the intensity value of the reference image B is known *a-priori*. Images

are registered spatially in such a manner that a maximum amount of information content (bits per voxel) can be reduced to restore the floating image given the reference image. It also implies that there is a maximum reduction of the amount of "choice" involved to predict the voxel intensity values on A′ by a measurement of voxel intensity values performed on B. When images are registered, the intensity value of a voxel in B is a good predictor of the intensity of the corresponding voxel in A′. The purpose of the registration algorithm is to find a transformation that maximizes MI between the transformed floating image $T_\theta A$ and the reference image B. The outcome of the registration algorithm using MI is given by $T_\theta = \arg\max_{T} MI(T_\theta A, B)$. Entropies of the individual images play an important role in registration problems since entropy depends on the size and information content of the overlap between the images. One consequence is that entropy of the images for registration, computed from the histogram, may change as a function of transformation. If the individual entropies are not taken into account, background voxels will bias the registration so that the images will be completely misaligned, as shown in the 2D example in Figure 3.9.

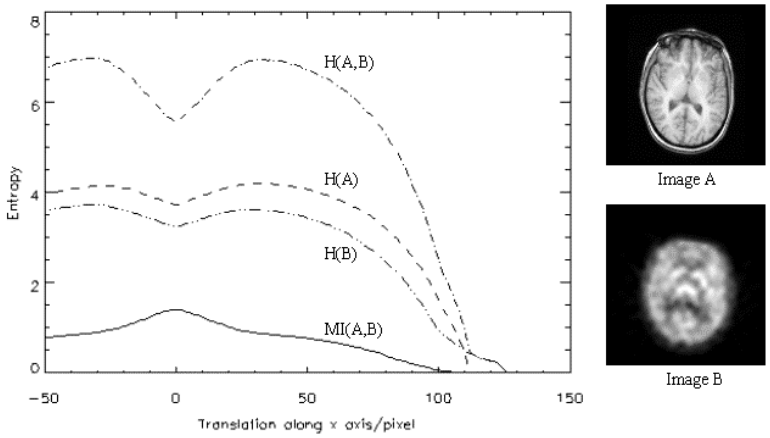


Figure 3.9: Entropies and mutual information of MR and SPET images are plotted as a function of translation along the x-axis. Each image has 128×128 pixels and 256 intensity bins. The measures are computed, with full sampling, in the overlap between the images. Entropies H(A,B), H(A) and H(B) are sensitive to translation. The global minimum of all entropies occurs beyond 128 pixels where the overlap between the images vanishes. MI has a global maximum at the zero translation (true alignment). Joint entropy H(A,B) shows a local minimum at zero translation.

This helps to explain why joint entropy is inferior to mutual information as the optimization criterion for image registration [48,269]. The effect of including the non-overlapping region on entropies is shown in Figure 3.10. The entropy of the MR image (reference image A) is constant while the entropy of the SPET (floating image B)

decreased gradually as a function of translation along x-axis. The entropy of MR is constant while entropy of the transformed SPET varies with translation as the interpolated position is outside the image domain of the original SPET.

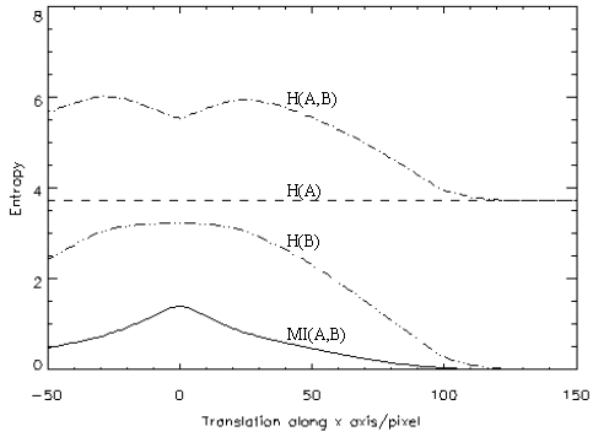


Figure 3.10: Entropies and mutual information (vertical axis) of MR (reference image) and SPET (floating image) are plotted as a function of translation in pixels along the x-axis. The measures are computed at full sampling and include the non-overlapping region. Entropy of MR denoted by H(A) is independent of translation while the entropy of SPET varies with the translation. The position of zero translation corresponds to the registered position found by MI(A,B).

Alternatively, mutual information can be expressed as a third frequently used definition in terms of the Kullback-Leibler distance [270], also known as cross-entropy employed in [271]. Given two random variables P and G, the K-L distance is defined as $d(P,G) = \sum_i P(i) \log_2 \frac{P(i)}{G(i)}$. Expanding Equation (3.10), the third definition of MI is obtained as the Equation (3.11)

$$\begin{aligned}
\mathrm{MI}(\mathrm{A},\mathrm{B}) &= -\sum_a \mathrm{P}(a)\log\mathrm{P}(a) - \sum_b \mathrm{P}(b)\log\mathrm{P}(b) + \sum_a\sum_b \mathrm{P}(a,b)\log\mathrm{P}(a,b) \\
&= -\sum_a\sum_b \mathrm{P}(a,b)\log\mathrm{P}(a) - \sum_a\sum_b \mathrm{P}(a,b)\log\mathrm{P}(b) + \sum_a\sum_b \mathrm{P}(a,b)\log\mathrm{P}(a,b) \\
&= -\sum_a\sum_b \mathrm{P}(a,b)\left\{\log\mathrm{P}(a) + \log\mathrm{P}(b)\right\} + \sum_a\sum_b \mathrm{P}(a,b)\log\mathrm{P}(a,b) \\
&= -\sum_a\sum_b \mathrm{P}(a,b)\log\frac{\mathrm{P}(a,b)}{\mathrm{P}(a)\mathrm{P}(b)} \\
&= d\big(\mathrm{P}(a,b), \mathrm{P}(a)\mathrm{P}(b)\big) \qquad (3.11)
\end{aligned}$$

The K-L distance compares two probabilities P(a,b) and P(a)P(b). If the occurrence or non-occurrence of intensity a in A has no effect on the occurrence of intensity b at the corresponding voxel in B, then A and B are said to be *independent*. The probability of occurrence of intensity in one image is the same whether or not the intensity in the other

image is known. Symbolically, if A and B are independent, then $P(a \mid b) = P(a)$ and $P(b \mid a) = P(b)$. Then Equation (3.6) becomes $P(a,b) = P(a)\,P(b)$. With reference to Figure 3.8, this corresponds to zero overlap between the entropy of A and B. From (Eq. 3.11), mutual information of A and B is zero. When measurement on A does nothing to reduce uncertainty about B, these relations hold: $H(A \mid B) = H(A)$, $H(B \mid A) = H(B)$ and $H(A,B) = H(A) + H(B)$ when $MI(A,B) = 0$ by virtual of independency between A and B. At the other extreme, A and B are *perfectly dependent*. We can predict one perfectly given the other and it follows that $H(A \mid B) = H(B \mid A) = 0$, and $MI(A,B) = H(A) = H(B) = H(A,B)$. Thus the entropies of A and B are the same.

3.4.4 Normalized mutual information

Recall that the total number of voxels used to normalize the joint histogram varies with the size of the overlap between the two images, and this changes with the transformation. Consider the case where the transformation maps some voxels in one *image* to fall outside the domain of the other image. The joint entropy H(A,B) and the marginal entropy of each image A and B all vary with the size of the overlap. Thus MI-based image registration can fail in such situations. To reduce the dependence of mutual information on the size of overlap, Studholm *et al.* [180] proposed a normalized version given by

$$NMI(A,B) = \frac{H(A) + H(B)}{H(A,B)}. \qquad (3.12)$$

They argued that when the marginal entropies H(A) and H(B) increase faster than the joint entropy H(A,B), mutual information MI(A,B) increases with misregistration as a result of a decrease of overlapping regions. To illustrate the dependence of entropies on the overlapping region, the marginal entropies H(A) and H(B), and joint entropy H(A,B), mutual information MI(A,B) and normalized mutual information NMI(A,B) were evaluated directly for a pair of 2D MR and SPET images with different rotational misalignments (Figure 3.11). The graphs show that NMI was able to predict the *true* registered position at zero rotation while MI gave a false maximum around ±20° rotations in axial direction. The plot on the left also show that H(A,B) increases faster than H(A) and H(B). It explains why NMI gives better registration than MI for this particular case.

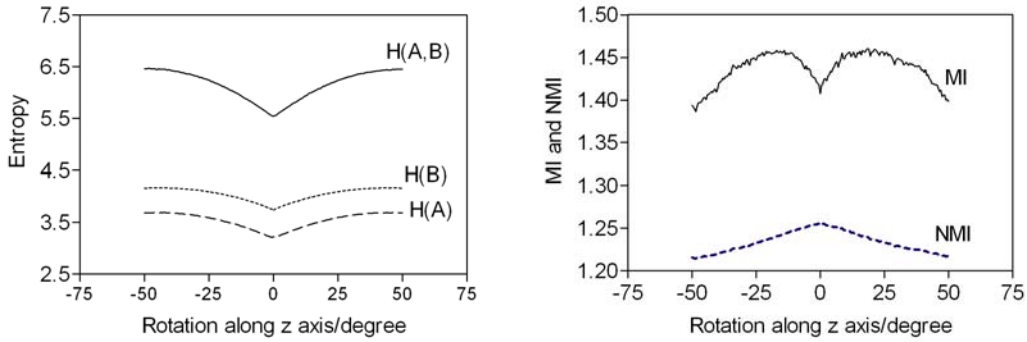


Figure 3.11: The left plot shows the effect of reducing overlapping region on the marginal entropy and joint entropy for SPET (image A, 128×128) and MR (image B, 128×128). These entropy measures are computed by using full sampling and plotted as a function of in-plane rotation in the overlap region between the images. The position of zero rotation is the registered position. Increase of marginal entropy and joint entropy as a function of rotation is found. The plot on the right shows the mutual information (MI) and normalized mutual information (NMI) (at full sampling) plotted as a function of in-plane rotation. The position of zero rotation is the registered position. On each side of zero, a local maximum is observed, which may lead to failure of registration using MI. The NMI is maximum at the position of zero rotation (registered position).

3.5 Variance-based algorithms

3.5.1 Image variance

In Section 3.1, intensity values of images are introduced as random variables. The joint intensity histogram was found to be an effective method of picturing a relationship between images. Beside entropy, classical statistical quantities such as the mean and variance were suggested by a number of research groups [131,172,175] to describe the joint histogram in a quantitative way. Again, let the intensity $a = \mathrm{A}(x)$, at each voxel in an image, be a random variable. They are all independent. Suppose a set of possible intensity values $a \in \{0,1,2,..,M-1\}$ whose probabilities of occurrence are known. Let the mean intensity μ be, by definition, the *expected* intensity of a sufficiently large image

$$\mu = \mathrm{E}_a[a] = \sum_a a\,\mathrm{P}(a) \qquad (3.13)$$

The dispersion of the joint histogram changes as a function of transformation and, therefore, it may differ greatly. To describe how sample intensity of the image is spread, the variance $\sigma^2(a)$ of the sample data is used to measure how far on average the observations will be from the sample mean. It is defined as the squared deviation on average of observations from the sample mean μ and is given by

$$\sigma^2(a) = \sum_a \mathrm{P}(a)\,(a-\mu)^2 = \mathrm{E}_a[(a-\mu)^2] \qquad (3.14)$$

When the sample intensity of a region is composed of a single tissue type, the variance of intensity values is expected to be small. To apply the statistical concept of variance for image registration, three assumptions have to be made.

First, anatomical structures can be delineated by virtue of tissues. The way intensity is presented in anatomical and functional images depends on the physics of the imaging process. Excellent soft tissue contrast and high signal-to-noise ratio are found for anatomical images such as CT and MR images. In SPET and PET images, voxel intensity values are frequently corrupted by noise, finite resolution and partial volume effect (presence of multiple tissues within a voxel) causing inhomogeneity of intensity within functional structures. Second, an image can be partitioned into subimages according to voxel intensity values. Voxel intensity values and volume elements (voxels) of tissue are related as a result of image acquisition. Thus the intensity value can be used to partition each voxel into one of M different tissue classes, $a \in \{0,1,...,M-1\}$. Third, corresponding anatomical structures are spatially aligned when images are registered. Together with the first assumption, this implies that a region of uniform intensity in one image will correspond to a region of uniform intensity in the other image when they are registered. Under these assumptions, image registration may be formulated as a process that maximizes the uniformity of intensity in one image given the knowledge of intensity of the other image. This is equivalent to an optimization algorithm that seeks to minimize variance of intensity in A when intensity in B is known.

3.5.2 Conditional variance

Given two images A and B, the search for the correct alignment can be cast as a variance minimization problem. As suggested by Viola [117], when the probability of a random variable "*a*" drawn from a distribution is high, the occurrence of "*a*" has smaller uncertainty (low entropy). A high and narrow peak corresponds to a small variance with low entropy where the observations will not be far away from the sample mean. It follows that a marginal probability distribution of a joint histogram with a narrow peak has low uncertainty because most of the measurements will fall in the

region where the density in PDF is high. A broad marginal distribution has a high uncertainty because the sample is spread out with a large variance, which corresponds to a high entropy and low density in PDF. This permits an interpretation of the uncertainty (entropy) and variance of the joint histogram in terms of probability. If A is dependent on B, knowledge of image B should reduce the variance (uncertainty) of image A [265]. It is useful to define a conditional expectation, which gives the mean intensity value of image A given the intensity value of image B. In multimodality registration, image B is generally the one with the better resolution. Intuitively, the conditional variance gives a measure of the average intensity for an induced region in A which is partitioned by the intensity "*b*" in B. The induced sub-region $S_b = \{x \in \Omega_A \cap \Omega_B : B(x) = b\}$ refers to a subimage S_b of A defined by the intensity "*b*" of image B, where Ω_A and Ω_B is the domain of image A and B respectively. The partition splits the image into disjoint subimages whose union is the image. The conditional expectation $E_a[a|b]$ denotes a mean intensity of an induced sub-region in A while $E_a[a]$ denotes the mean of intensity values in A as a whole. Since $E_a[a]$ and $E_a[a|b]$ originate from the joint histogram, they depend on the overlap region and, consequently, on the transformation. Let the mean intensity of induced region in A by the intensity *b* in the image B be μ_b, then

$$E_a[a|b] = \sum_a a\, P(a|b) = \mu_b \tag{3.15}$$

The variance of intensity in each induced region in A corresponding to a given region of intensity "*b*" in B and is denoted as $\sigma^2(a \mid b)$, where $a, b \in \{0,1,2,...,M-1\}$. The conditional variance evaluates the homogeneity of intensity within each induced region in A. It is computed from the conditional probability distribution derived from the joint intensity histogram for each *b*, defined by

$$\begin{aligned} \sigma^2(a|b) &= \sum_a P(a|b)\,(a - E_a[a|b])^2 \\ &= E_a[(a - E[a \mid b])^2] \end{aligned} \tag{3.16}$$

The variance $\sigma^2(a \mid b)$ per induced segmentation in A via the intensity "*b*" of B is weighted by its probability of occurrence P(*b*) in B as shown in the Equation (3.17). The higher the frequency of the intensity in B, the greater the weight is assigned to the variance. The expected variance $E_b[\sigma^2(a \mid b)]$, which is known as *conditional variance*,

gives the average amount of dispersion of the joint histogram of A given the knowledge of B. In other words, $\mathrm{E}_b[\sigma^2(a \mid b)]$ measures the uncertainty of A when B is known by means of variance rather than relying on entropy. Recall that a high and narrow peak histogram gives a small variance corresponding to a low entropy (small uncertainty) probability distribution. The expected value of the conditional variance is defined in terms of the marginal and conditional probabilities,

$$\begin{aligned}\mathrm{E}_b[\sigma^2(a \mid b)] &= \sum_b \mathrm{P}(b)\,\sigma^2(a \mid b) \\ &= \sum_b \sum_a \mathrm{P}(b)\mathrm{P}(a \mid b)(a - \mathrm{E}_a[a \mid b])^2 \\ &= \sum_b \sum_a \mathrm{P}(a,b)(a - \mathrm{E}_a[a \mid b])^2 \end{aligned} \qquad (3.17)$$

Using Equation (3.15), $\mathrm{E}_b[\sigma^2(a \mid b)]$ can be expressed more succinctly as

$$\mathrm{E}_b[\sigma^2(a \mid b)] = \mathrm{E}_{ba}[(a - \mu_b)^2]\,. \qquad (3.18)$$

A transformation where the images become disjoint will minimize the conditional variance although the images are grossly misaligned. To avoid a minimum when the overlap is small, one approach is to normalize the conditional variance by the image variance [264]. The normalized conditional variance (NCV) is given by

$$\frac{\mathrm{E}_{ba}[(a - \mathrm{E}_a[a \mid b])^2]}{\sigma^2(a)} = \frac{\sum_b \sum_a \mathrm{P}(a,b)\,(a - \mu_b)^2}{\sum_a \mathrm{P}(a)\,(a-\mu)^2}\,, \qquad (3.19)$$

where $\mu_b = \sum_a a\,\mathrm{P}(a|b)$, $\mu = \sum_a a\,\mathrm{P}(a)$, $\mathrm{P}(a) = \sum_b \mathrm{P}(a,b)$ and $\mathrm{P}(b) = \sum_a \mathrm{P}(a,b)$ are all computed from the joint histogram. In this approach, image registration could be regarded as a least square optimization weighted by the occurrence of intensity pairs in the joint histogram. Note that the conditional variance is asymmetrical,

$$\sum_b \sum_a \mathrm{P}(a,b)\,(a - \mu_b)^2 \neq \sum_b \sum_a \mathrm{P}(a,b)\,(b - \mu_a)^2\,.$$

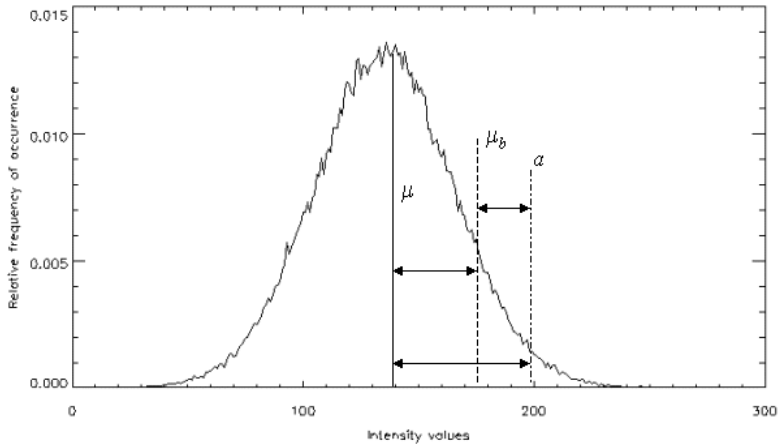


Figure 3.12: An illustration of the relationship between $\sigma^2(a)$, μ, μ_b and "a" is given by Equation (3.20). The symbol μ denotes the overall average intensity in A. μ_b gives the average intensity of an induced region in A which corresponds to a region of the intensity "b" in B. $\sigma^2(a)$ is the sum of the square deviation of intensity "a" from the mean intensity in A. Conceptually, the difference between μ_b and μ represents the reduction of uncertainty of a sub-region of A given the knowledge of B. The difference between "a" and μ_b represents the uncertainty in predicting the intensity in the induced sub-region of A given the intensity "b" in B. The relative frequency of occurrence of each intensity (vertical axis) is plotted against the intensity value (horizontal axis).

3.5.3 Correlation ratio

If A and B are dependent, the variance of intensity in A can be explained by the knowledge of B,

$$\begin{aligned}\sigma^2(\mathrm{a}) &= \sum_a \mathrm{P}(a)\,(\mathrm{a} - \mathrm{E_a}[a])^2 \\ &= \sum_a \sum_b \mathrm{P}(a,b)(a - \mathrm{E}_a[a])^2 \\ &= \mathrm{E}_{ab}[(a - \mathrm{E}_a[a])^2] \\ &= \mathrm{E}_{ab}[(a - \mu)^2].\end{aligned}$$

If $\mathrm{E}_a[a \mid b]$, which gives a mean intensity of a sub-region of A corresponding to the sub-region of intensity "b" of B, is inserted inside the curly brackets, $\sigma^2(a)$ can be given as

$$\begin{aligned}\sigma^2(a) &= \mathrm{E}_{ab}\{(\mathrm{E}_a[a \mid b] - \mathrm{E}_a[a \mid b] + a - \mathrm{E}_a[a])^2\} \\ &= \mathrm{E}_{ab}\{(\mu_b - \mu + a - \mu_b)^2\} \qquad (3.20)\end{aligned}$$

A schematic diagram showing the relation of $\sigma^2(a)$, μ, μ_b and "a" is given in Figure 3.12. When P(a) and P(b) are independent, $\mathrm{P}(a|b) = \mathrm{P}(a)$ and $\mathrm{E}_a[a \mid b] = \mathrm{E}_a[a]$. P($a$) is the prior distribution and P($a|b$) is the posterior distribution. Unless A and B are independent, the posterior P($a|b$) gives some information about P(a) and the intensity of

"*b*" in B is useful to predict the intensity in A. The difference between the most probable values of $\mathrm{E}_a[a \mid b]$ and $\mathrm{E}_a[a]$ give a Bayesian information gain which reduces the uncertainty about A given the knowledge of B. Note that if $\mathrm{E}_a[a] = \mathrm{E}_b[\mathrm{E}_a[a \mid b]]$, the conditional expectation is unbiased with respect to reference image intensity "*b*". This can be shown as follows:

$$\begin{aligned}
\mathrm{E}_b[\mathrm{E}_a[a \mid b]] &= \mathrm{E}_b[\sum_a a\, \mathrm{P}(a|b)] \\
&= \sum_b p(b) \sum_a aP(a \mid b) \\
&= \sum_a \sum_b aP(a \mid b)P(b) \\
&= \sum_a \sum_b aP(a,b) \\
&= \sum_a aP(a) \\
&= \mathrm{E}_a[a]
\end{aligned}$$

Expanding Equation (3.20), we have

$$\begin{aligned}
\sigma^2(a) = \mathrm{E}_{ba}[&(\mathrm{E}_a[a \mid b] - \mathrm{E}_b[\mathrm{E}_a[a \mid b]])^2 + (a - \mathrm{E}_a[a \mid b])^2 \\
&+ 2(\mathrm{E}_a[a \mid b] - \mathrm{E}_a[a])\,(a - \mathrm{E}_a[a \mid b])]
\end{aligned}.$$

Since $(\mathrm{E}_a[a \mid b] - \mathrm{E}_a[a])\,(a - \mathrm{E}_a[a|b]) = 0$ and $\mu_b = \sum_a a\, \mathrm{P}(a|b) = \mathrm{E}_a[a|b]$, $\sigma^2(a)$ can be expressed as

$$\sigma^2(a) = \sigma^2(\mu_b) + \mathrm{E}_b[\sigma^2(a \mid b)]. \qquad (3.21)$$

For more details the reader may refer to the reference [181]. It is interesting to note here that the variance of A can be split into two parts. $\sigma^2(\mu_b)$ is the part of A that can be predicted by B whereas $\mathrm{E}_b[\sigma^2(a \mid b)]$ is the part of A that is independent of B. The correlation ratio CR is defined by [181]:

$$\mathrm{CR} = \frac{\sigma^2(\mu_b)}{\sigma^2(a)} = 1 - \frac{\mathrm{E}_b[\sigma^2(a \mid b)]}{\sigma^2(a)} \qquad (3.22)$$

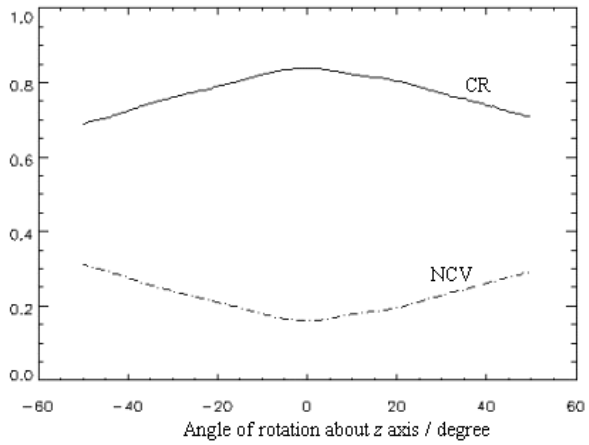


Figure 3.13: Correlation ratio CR and normalized conditional variance NCV (at full sampling) for MR (128×128) as the reference image and SPET, denoted the floating image, (128×128) is plotted as a function of in-plane rotation. CR is maximum while NCV is minimum at zero rotation (true alignment).

For multimodality registration, a linear dependence between images is unlikely. Unlike the correlation coefficient [272] that computes linear relationship between voxel intensities, CR relies on the variance to measure the functional dependence between two images and it is therefore useful for multimodality registration. CR equals 1 when two images are perfectly dependent and is zero when the images are independent. CR is related to the conditional variance (Equation 3.19) by a constant term. Maximizing CR is equivalent to minimizing the normalized conditional variance NCV. Unlike mutual information, CR is not symmetrical and depends on which image is used to predict the other. Figure 3.13 shows the resulting CR and NCV for rotational displacement of up to ±50° about the z-axis. Maximization of CR seeks a transformation where the normalized conditional variance is minimized. To compute CR from the joint histogram, the following equation can be used.

$$\mathrm{CR} = 1 - \frac{\sum_b \sum_a \mathrm{P}(a,b)\,(a - \mu_b)^2}{\sum_a \mathrm{P}(a)\,(a - \mu)^2} \,. \qquad (3.23)$$

The correlation ratio is closely related to an earlier measure proposed by Woods *et al*. [172] and described in detail in Subsection 2.2.2. Their similarity measure (Equation 2.4) can be rewritten, using the notation of this chapter, as $\mathrm{E}_b\left(\sigma(a \mid b) / \mu_b\right)$ where $\sigma^2(a \mid b)$ is the conditional variance (Equation 3.16) and μ_b refers to the mean intensity of an induced region in PET. This algorithm partitions the MR image (B) based on the

MR intensity, and maximizes the uniformity of intensity within each induced partition in the PET image (A) for registration. Woods *et al.* [172] minimize the standard deviation of the PET intensity normalized by its mean whereas Roche *et al.* [181,265] minimize the conditional variance of intensity normalized by $\sigma^2(a)$ (Equation 3.22). The Woods' technique has been widely used for MR-PET registration [48] but its use in other modalities is not documented. A drawback of the Woods' criterion is that it requires an extra step to edit out the non-brain regions (i.e. skin, skull, and dura) from MR, otherwise bias might occur in $P(a|b)$ which may lead to local extrema. This can be done manually or automatically [273]. Without this step, the performance of Woods' algorithm in MR T_1-T_2 and PET-MR registration was consistently better than MI but less accurate than CR when sparse sampling was used [265].

3.5.4 Extension - Symmetric correlation ratio

A symmetrical correlation ratio is a measure independent of which image is used to estimate the other. A simple approach is to evaluate the normalized conditional variance as the sum of two CRs [182]

$$\mathrm{SCR} = \mathrm{CR}(\mathrm{A},\mathrm{B}) + \mathrm{CR}(\mathrm{B},\mathrm{A}), \qquad (3.24)$$

where CR is defined in Equation (3.23). The mean of CR(A,B) and CR(B,A) would be an alternative approach to retain symmetric property of the similarity measure by CR. The symmetric correlation ratio is independent of the size of the overlap between the images A and B. Here $\sigma^2(a)$ and $\sigma^2(b)$, estimated from the joint histogram, are the variances over the overlap region defined by the relative transformation of images A and B. The joint probability $P(a,b)$ and the marginal probabilities $P(a)$ and $P(b)$, as well as the conditional probabilities $P(a|b)$ and $P(b|a)$ are computed from the normalized joint and the marginal intensity histograms, respectively.

3.6 Conclusion

Registration of different image modalities is the preliminary and mandatory step to combine anatomical and functional information. The integration of multiple complementary data into a common reference allows a more comprehensive analysis of patient treatment, which is not available while looking at images from a single modality. For the purpose of image registration, 3D medical images are taken as comprising of voxels and each voxel is treated as a random variable. An important consequence of

treating images as random variables is that statistical measures can be used to investigate the image registration problem. In this chapter a theoretical framework of the joint intensity based registration method has been presented, which is coarsely divided into approaches based on information theory versus statistical measures. In the first category, registration relies on the entropy; in the latter, registration relies on mean and variance. The use of a two-dimensional plot showing the combinations of intensity values in each of the two images for all the corresponding voxel pairs is found to be useful for image registration. Given a joint intensity histogram of two images, entropy measures information content, uncertainty and dispersion of the intensity pairs of these two images. In connected or dis-jointed sub-regions of an image, mean measures the central tendency of intensities and variance measures the dispersion around the mean. These sub-regions are based on the clustering in the joint intensity histogram. The underlying process of how misregistration influences the pattern of the joint intensity histogram, the joint probability distribution and the marginal probability distribution is subtle and difficult to understand. Over the past few years a large amount of image registration has been applied based on the use of mutual information, normalized mutual information and correlation ratio for multimodality images registration. From the diversity of studies reported in literature, it is clear that these similarity measures are generally applicable for a wide range of multimodality image registration without preprocessing, user initialization or parameter tuning. From the conclusion of certain comparison studies that mutual information is not a cure for all image registration; adaptation (e.g. normalized mutual information [180]) and alternative approaches (e.g. correlation ratio [264,181]) are suggested by a number of research groups. A modified correlation ratio is proposed to retain the symmetric property of this similarity measure. To assess the feasibility of using this proposed similarity measure for multimodality image registration; the accuracy of the technique needs to be evaluated both by phantom experiments and on patient data. The implementation, the effect of subsampling, intensity rebinning, interpolation, optimization and its accuracy will be discussed in the next chapter.

Chapter 4

Comparison of SCR with other joint histogram techniques

4.1 Introduction

Several different approaches have been used in the study of multi-modality image registration based on the use of the joint intensity histogram technique. Similarity measures, including mutual information (MI), normalized mutual information (NMI) and correlation ratio (CR), have been demonstrated to be applicable for image registration. Thurfjell *et al.* [258] demonstrated that intensity rebinning and subsampling improve the efficiency of SPET-MR registration in brain, while maintaining registration accuracy. The improvement was demonstrated for several cost functions including CR. However, CR is not symmetric in that the outcome changes when images are swapped. The symmetrical correlation ratio (SCR) is proposed as an alternative version of CR to improve its symmetrical property [182]. Broadly speaking, MI and NMI are entropy-based algorithms while CR and SCR are variance-based methods. These very different approaches may have different registration performances depending on experimental techniques, pre-registration processing and/or imaging modalities. The experiments performed in this chapter concern the practical aspects of implementation, such as subsampling schemes, intensity rebinning, interpolation, multi-level sampling and optimization.

The aim of this chapter is to test and validate the proposed symmetrical correlation ratio (SCR) for matching T_1-T_2 MR and MR-SPET brain scans. It also presents a comparison of SCR with MI, NMI and CR on registration accuracy achievable using various parameters. The evaluation consists of three studies to assess the performance of SCR. The first experiment involved a clinical data set of T_1 and T_2 MR image pairs with simulated misalignment or rigid transformation to study the effect of intensity rebinning and various subsampling techniques on registration accuracy. Determining the Euclidean distance of vertices of a bounding box before and after registration assesses the matching error. In the final experiment, accuracy in clinical MR and SPET images is measured. Quantification of the registration error is based on the spatial distance of a set of external fiducial markers between MR and SPET images after registration.

4.2 Measures of performance

To assess the registration algorithm, a numbers of approaches have been reported. Interested readers may refer to Chapter 2 (Subsection 2.6.2) for a more detailed discussion on metrics for evaluation of registration methods. Registration accuracy gives a direct interpretation of registration performance if it is measured by Euclidean distance. It is generally determined from the spatial distance between corresponding pairs of points on a floating image and a reference image after registration. In the present study, two approaches are used to determine accuracy after registration: displacement of selected points and displacement of fiducial markers.

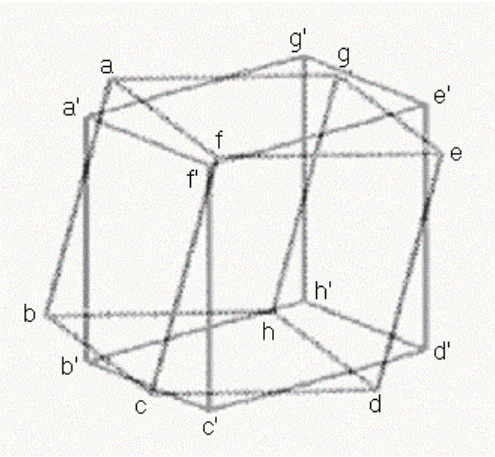


Figure 4.1: Error is computed from the distance between the corresponding corners (a,b,c,d,e,f,g,h) of a bounding box after registration. After registration the corners became (a′,b′,c′,d′,e′,f′,g′,h′).

4.2.1 Displacement of selected points

When clinical image data is used, it is difficult to evaluate registration accuracy because ground truth is, in general, unknown. To resolve this problem, a known transformation may be applied to one image and a second image is then registered to it. This allows an estimate of the spatial discrepancy of selected points after registration to be made. In one approach, selected points are the vertices of a bounding box around a pair of registered MR brain images, such as proton-weighted (image A) and T_2-weighted (image B) MR. To assess registration accuracy, the reference image B is relocated by a known random transformation. The floating image A is registered to B to recover the transformation. The registration error is determined by the average Euclidean displacement between corresponding points over all the eight vertices of the bounding box after registration [179,242,258]. Figure 4.1 illustrates the distance between corresponding vertices. $\theta_{\text{app}}(x)$ denotes the transformation applied to the reference

image B, and $\theta_{\text{rec}}(x)$ refers to the recovered transformation after A is registered to B. The displacement of each vertex $\left\|\theta_{\text{rec}}(x_i) - \theta_{\text{app}}(x_i)\right\|$ was computed and averaged over all vertices to yield the displacement error, $\text{DE} = \frac{1}{8}\sum_{i=1}^{8}\left\|\theta_{\text{rec}}(x_i) - \theta_{\text{app}}(x_i)\right\|$, to measure registration accuracy.

4.2.2 Displacement of fiducial markers

The use of simulated transformation to assess registration accuracy is convenient because it permits controlled evaluation over a wide range of conditions for image acquisition. However, it is not a *real* clinical misalignment studies to evaluate *actual* registration accuracy. In the absence of salient features, registration error is best characterised by the distance between external fiducial markers after registration [75,122,123,236]. After registration the averaged Euclidean displacements over all markers for each subject gives the fiducial registration error (FRE). The FRE is defined as $\text{FRE} = \frac{1}{\text{n}}\sum_{i=1}^{\text{n}} \| \mathbf{Q}\, x_i - y_i \|$, where n denotes the number of fiducial markers, x_i and y_i is position vectors of the corresponding marker *i* on the floating image and the reference image respectively. **Q** denotes the matrix representing the recovered transformation.

4.3 Experimental techniques

The problem of image registration is to find the best spatial correspondence between two images. The stationary image to be registered is referred to as the reference image (B) and the image that is registered is denoted as the floating image (A). The reference image center is taken as the origin of the coordinate system for the transformation. All image alignment is restricted to the rigid-body type in this chapter. The registration parameters θ are represented by a 6-dimensional vector (φ_x, φ_y, φ_z, t_x, t_y, t_z) where φ_x, φ_y, φ_z, are rotation angles in degrees about the x-, y- and z-axis respectively, and t_x, t_y, t_z are translation offsets in mm along the x-, y- and z-axis, respectively. When a transformation T_θ is applied to the floating image, a transformed floating image $T_\theta A$ is generated. For the purpose of registration, a transformation that minimizes a negated similarity measure Φ of $T_\theta A$ and B need to be estimated. Thus the registration process is to find T_θ such that

$$\hat{\mathrm{T}}_\theta = \arg\min_{\mathrm{T}} [\Phi(\mathrm{T}_\theta \mathrm{A}, \mathrm{B})], \tag{4.1}$$

evaluated over the region of overlap of the two images. Two images A (floating image) and B (reference image) are defined over a 3D spatial coordinate system. The intensity pair $(a,b) = (\mathrm{A}'(x), \mathrm{B}(x))$ is defined over the overlap region $x \in \Omega_{AB}$ of the two images, where a is the interpolated voxel intensity in the transformed floating image A′ and b is the voxel intensity in the reference image B. The intensity a of A′ is related to A via a spatial transformation T and an interpolator I_{AB}. In registration, transformation T maps A to A′ where $\mathrm{A}'(x) = \mathrm{T}_\theta \mathrm{A}(x) = I_{AB}\{\mathrm{A}(\mathbf{Q}\,x)\}$. $\mathrm{A}'(x)$ may not coincide with the grid point of Ω_{A} and local interpolation is needed to compute the intensity of A′ at x. See Section 3.2 for details. For a given transformation T_θ, the similarity measure Φ is computed via a joint histogram created by plotting a point (a,b) for every pair of corresponding voxels in $x \in \Omega_{AB}$.

4.3.1 Subsampling and multi-level sampling

A typical rigid-body registration could involve several hundred iterations to evaluate a similarity measure, depending on the image size, optimization algorithm, parameteric space dimension and interpolation technique. For example, image pairs of size 128×128×128 involve 2^{21} interpolations to update the joint histogram for each degree of freedom in the parameteric space making the registration algorithm computationally expensive. An effective tactic is to use a multi-resolution approach, subsampling scheme and/or an intensity rebinning to speed up the process [218,225,244,258,269,274,225,271]. To construct the joint intensity histogram, different sampling rates can be applied to each axis. Given sampling rates n_x, n_y and n_z along the respective axes, the total number of voxels is reduced by a factor of $n_x n_y n_z$. Since the computational cost of sampling images (T_θA and B) and the calculation of the similarity measure varies linearly with the number of samples, this subsampling method accelerates the registration process by a factor of $n_x n_y n_z$. Furthermore, in a multi-resolution and multi-levels approach, coarsely sampled images are registered first and the result serves as the starting point for the registration of more finely sampled images to fine tune the alignment.

Two methods have been previously proposed to perform subsampling for image registration. The first method creates a stack of progressively blurred images, e.g. by averaging neighbourhood voxels, prior to subsampling. It is done by down sampling (subsampling) a blurred version of the images $T_\theta A$ and B at low resolution to construct the joint histogram and then progressively refining the sampling interval at increasing resolution of the images. This strategy has been employed to register intra- and inter-modality images in [175,232,244,258,275]. One possible side effect of this approach is that the smoothing of images may change the joint histogram in such a way that the high frequency component information of the image will appear different at a lower resolution. In the second approach, a multi-level sampling scheme is applied to the transformed floating image $T_\theta A$ and the reference image B at full resolution using integral subsampling factors to construct the joint histogram. In this case, the problem of the dispersal of image information from high frequencies to low frequencies in the joint intensity histogram can be reduced because the image intensity values will not change as a result of blurring.

To do the coarse sampling, one may transform the floating image A as a whole and then subsample $T_\theta A$ and B at the integral subsample grid points defined in the reference image coordinates. Without reducing the total number of voxels used for transformation, this method is computationally expensive. The problem can be circumvented by only using a fraction of the voxels (grid points) pre-determined along the three orthogonal axes via subsampling factors respectively. The grids are then transformed, and interpolation is used to reconstruct the intensity values in $T_\theta A$ corresponding to the transformed voxel locations at the center of these grids.

4.3.2 Interpolation

Evaluation of the criterion of registration as defined in Equation (4.1) involves resampling of the transformed SPET and computation of SCR. The computational cost of interpolation is an important consideration as it is required many times during the iterative process for image registration. When a transformation is applied to the floating image A, $T_\theta A$ is resampled implicitly according to the parameters of the spatial transformation model. To speed up the registration process, one may use a lookup table to infer the precomputed weights of the interpolation kernel without repeating the

computation to speed up the registration [276]. Due to the discrete nature of the floating image, the transformed position of a voxel in $T_\theta A$ will not coincide exactly with a voxel position in A. Intensity resampling is usually required to reproduce $T_\theta A$ by intensity interpolation at grid points $\mathbf{Q}x$ of the transformed image. Although higher order interpolators approximate the ideal sinc interpolation more closely, trilinear interpolation is used for its speed. One difficulty in performing interpolation is that parts of the floating image may fall outside the grid of the floating image. A predefined intensity value (other than zero) may be used to label voxels outside the floating image to solve this problem. This intensity value is then ignored in computing the joint histogram.

4.3.3 Histogram and intensity rebinning

After the floating image has been transformed and resampled, the image intensities of the voxel pairs are used to construct the joint intensity histogram from which the marginal, conditional and joint probability density functions (PDF's) are estimated from the normalized joint histogram. The voxel intensity of the two images is quantized into a fixed number of discrete bins (referred to as intensity bins hereinafter). The total number of bins is an important parameter of the joint-intensity histogram-based registration. The reference image and the floating image can have different numbers of intensity bins. To estimate the joint histogram the frequency of occurrence of intensity pairs is plotted in a 2D array of size $M_f \times M_r$ where M_f and M_r correspond to the number of intensity bins of the floating image and the reference image, respectively. Rebinning is a requantisation of the intensity values of the joint intensity histogram. Reducing the number of bins is equivalent to reducing the detail of the joint histogram. Considering 8-bit encoded images, it is done implicitly by rescaling all intensity values of images linearly that lie in the range $\{0,1,2,\ldots,255\}$ into the range $\{0,1,2,...,M-1\}$ where M is the intended number of intensity bins of the joint intensity histogram. Suppose the image size is 128×128×128 subsampled by a factor of 4 along each axis and $M_f = M_r = 256$. The sample data are reduced from 2^{21} to 2^{15} giving, on average, only 0.5 voxels per bin in the joint histogram. Thurfjell *et al.* [258] showed that reducing the number of bins yields smoother responses with fewer local minima. It was also demonstrated that a registration algorithm incorporating a coarse to fine sampling scheme with adaptive intensity rebinning was more computationally efficient without compromising

registration accuracy. By adaptive rebinning, the joint intensity histogram is progressively rescaled linearly in a coarse-to-fine manner.

4.3.4 Multi-level strategy

It has previously been shown that multi-level sampling and adaptive intensity rebinning is effective in reducing image registration time and avoids local minima [258]. In this approach, coarse images are registered first at a very low computational cost. The resultant registration parameters are the starting estimate for the finer-image registration at the next sampling level. In order not to change the intensity distribution of the images and hence the joint histogram, no image blurring is applied prior to registration. The schemes of [8,8,1|24], [8,4,1|32] and [4,4,1|64] are used for a three level registration. In this notation, $[n_1,n_2,n_3|M]$ denotes down sampling of the transformed floating image $T_\theta A$ and the reference image B by a factor of $[n_1,n_2,n_3]$ in the x-, y- and z-direction respectively, and M is the bin size of the intensity histogram. The registration process will take the difference of voxel sizes of images into consideration during computation to obtain the optimum registration.

4.3.5 Optimization

The optimal registration is usually found by iteratively minimizing the negated similarity measure (e.g. MI, NMI and SCR). The simplex algorithm described by Press *et al.* [127] is adopted in this work. Compared to gradient-based optimization techniques, the simplex algorithm is more efficient and simpler to implement because it does not require one-dimensional minimization or calculation of derivatives. Several independent studies have shown the simplex algorithm to be as robust as conjugate gradient and Levenberg-Marquardt algorithms for rigid-body registration [225,277]. The optimization starts its search with the centers of the two image volumes aligned. The simplex algorithm takes a series of steps to optimize the similarity. In the rigid-body registration, a 6 dimension (6D) parametric space corresponding to the six transformation parameters is employed. To search for the minimum of the similarity measure, a 6D hypercube is reflected, expanded and contracted in the parametric space. At each sampling level, termination occurs when the absolute change of the similarity value drops below 0.00001. A limit of 100 iterations was set as a secondary criterion to terminate the process at each level. Thus a maximum of 300 evaluations is performed for the 3-level registration.

4.3.6 Default implementation

The settings for the three-level subsampling scheme, trilinear interpolation and simplex optimization, when taken together are referred to as the default settings in the rest of this chapter unless otherwise specified.

4.4 Registration accuracy in clinical data with simulated misalignment

In this experiment, 3D anatomical image data sets were used to evaluate the inter-modality registration performance of the proposed similarity measure (SCR) compared to MI and NMI.

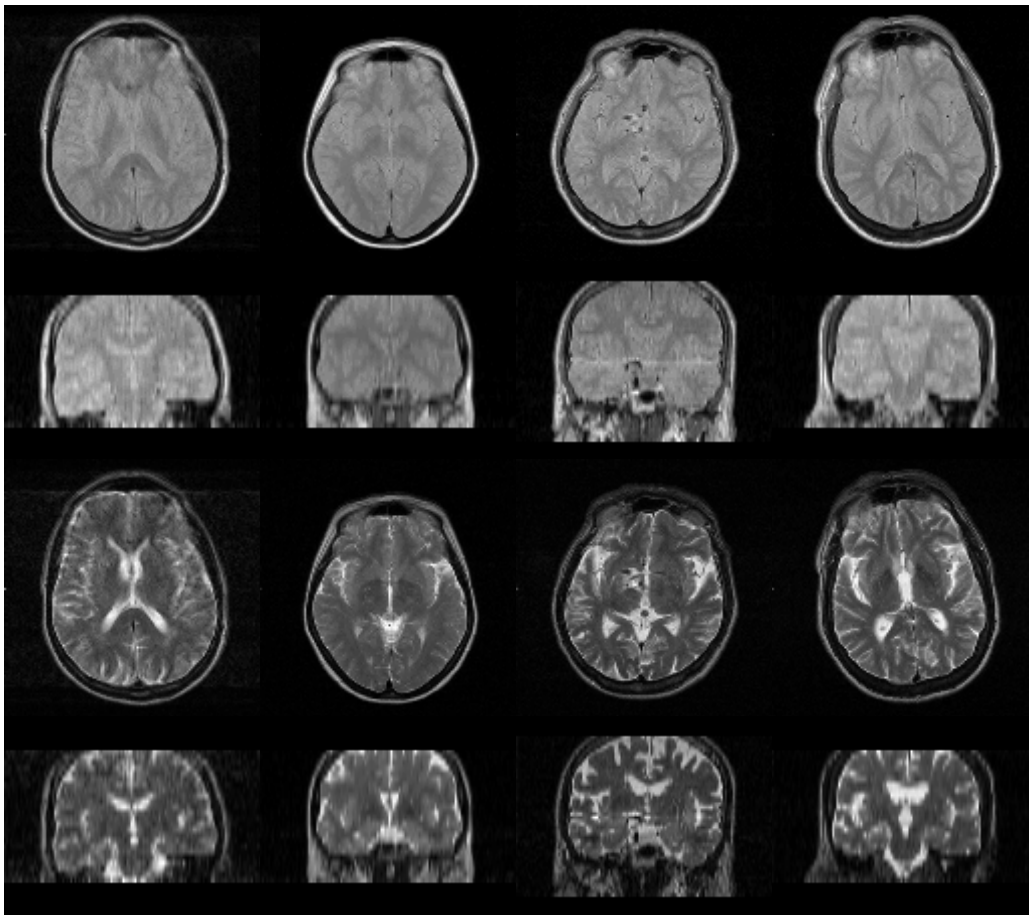


Figure 4.2: Examples of MR data used for evaluation. The diagrams show transaxial slices and coronal slices through four of the T_1 (upper block) and T_2 (lower block) MR images. Each image comprises $128\times128\times99$ voxels of size $(1.8\ \text{mm})^3$.

4.4.1 Experimental setup

Image data set

In order to assess the accuracy of registration for entropy-based algorithms (MI and NMI) and variance-based algorithms (CR1, CR2 and SCR), 10 independently acquired

MR T_1 and T_2 images of the brain were used. The original MR data set containing 256×256×24 voxels of size 0.9×0.9×5.0 mm^3 were converted to 128×128×99 cubic voxels of size 1.8 mm). Four of the MR images are shown in the Figure 4.2. Each image pair was registered during the acquisition as no misalignment was detected between the corresponding image data sets

Image misalignment

To obtain T_1 MR with a known difference in orientation as compared with the corresponding T_2 MR, the T_1 MR was repositioned using a rigid transformation. The transformed T_1 image is donated to as the reference image B, and the T_2 image is denoted as the floating image A. The parameters used for the transformation were generated in random order from a uniform distribution bounded by ±15° rotation about each axis and ±27 mm translation along each axis. The T_2 image was then registered to the transformed T_1 image using rigid registration. This process was repeated 3 times for each of the ten pairs of T_1 and the T_2 data giving a total of 30 registrations. The same experiments were performed for the entropy- and the variance-based similarity measures.

Experimental parameters

To assess the effectiveness of implementation, three different settings were used. These were: (1) the default setting (Subsection 4.3.5); (2) a modified default setting S1 and S2 in which adaptive intensity rebinning was discarded. The registration was repeated for 256 and 128 intensity bins independently, with S_1={[8,8,1|256], [4,8,1|256], [4,4,1|256]} and S_2={[8,8,1|128], [4,8,1|128], [4,4,1|128]}; (3) a modified setting in which multi-level strategy was abandoned. The number of intensity bins was fixed at 128 throughout the registration process. Three sampling schemes [2,2,2], [4,4,4] and [8,8,8] were used for image registration. When the registration algorithm proceeded through multiple sampling levels as in the cases of (1) and (2), simplex optimization [127] was used. A similar termination criterion as in Subsection 4.3.5 was applied to the optimization.

Preliminary investigation showed that 300 iterations were sufficient to ensure the registration to converge, as a further increase in number of iterations did not improve the registration. For the modified setting in (3), the registration process was terminated after 300 iterations or by a predefined threshold condition of absolute change of

similarity measure by 0.0001. To evaluate the registration accuracy a bounding box around the MR images was used (Subsection 4.2.1). Registration error was determined from the displacement of the corresponding vertices after registration. The average of the displacement error for all vertices, taken over all 30 registrations was referred to as the mean registration error. The median error, maximum and standard deviation of error were also determined to assess the registration performance. Two-tailed unpaired t-tests was used to compare mean registration error between similarity measures statistically, whenever possible.

4.4.2 Results and Discussion

Comparison of similarity measures - default implementation

The result of T_1-T_2 MR registration is given in Figure 4.3 (left). All measures performed similarly. This is confirmed by statistical test (one-way non-parametric ANOVA) and no significant differences (p<0.05) at 5% level were found among mean registration errors. The entropy- and the variance-based algorithms were all shown capable of registration accuracy comparable to the voxel size (1.8 mm) for anatomical image registration. West *et al.* [122] reported similar results for CT-MR registration based on mutual information.

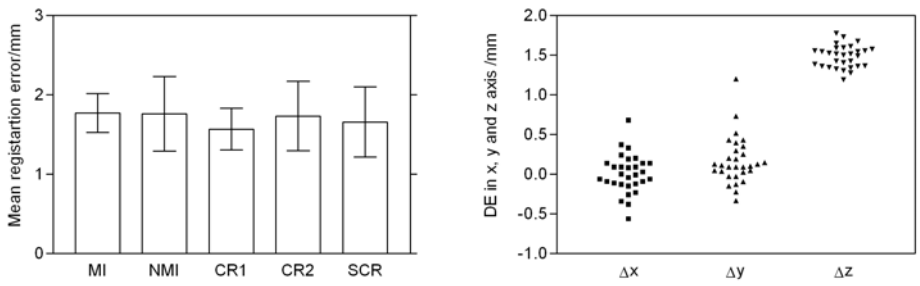


Figure 4.3: (Left): Mean registration error for MR T_1-T_2 registration for various similarity measures using sampling scheme of {[8,8,1|24], [8,4,1|32] and [4,4,1|64]}. Error bars show the standard deviation of each distribution. (Right): Displacement error averaged over eight vertices of the bounding box along x, y and z-direction for SCR registration.

A typical registration time was 56 seconds for SCR using a stand alone PC (666 MHz) with 192 Megabytes of random access memory. With full sampling, the time rose to 870 seconds so the default implementation accelerates the registration process by a factor of about 15. Summary of statistics of the registration is given in table 4.1 for each similarity measure with default implementation. Again, median and maximum

registration errors were all comparable among similarity measures. It is worth noting that the residual displacement error was not isotropic in the experiment. A typical in-plane residual displacement error was about 0.2 mm while residual displacement error along the axial direction was 1.5 mm as shown in Figure 4.3 (right). This may be due to the lower image resolution along the axial direction of the MR with the 1.5 mm error corresponding to 5 mm voxel depth and 0.2 mm error to 0.9 mm in-plane voxel size.

	MI	NMI	CR1	CR2	SCR
Mean ± SD (mm)	1.8 ±0.3	1.8 ±0.5	1.6±0.3	1.7 ±0.4	1.7±0.4
Median (mm)	1.7	1.6	1.6	1.6	1.6
Max (mm)	2.4	3.5	2.2	3.4	3.8

Table 4.1: Registration errors using default sampling scheme: [8,8,1|32], [4,8,1|48], [4,4,1|64]

Intensity rebinning

The results of registration using fixed numbers of bins for registration are given in the Figure 4.4. The scatter plots reveal sub-voxel accuracy for variance-based measures and NMI. Except for MI, the errors distribution for NMI, CR1, CR2 and SCR registration are almost identical among similarity measures, irrespective to 128 and 256 numbers of bins. The plots also show little, if any, effect of number of intensity bins on registration, and this is supported by the statistical test. The results of t-test (two-tailed unpaired) at 5% level show that the difference of mean registration error between modified default implementations S1, S2 and the adaptive rebinning schemes were not significant regardless of NMI, CR1, CR2 and SCR, with an exception in MI. Furthermore, typical registration accuracy for variance-based similarity measures were of the same order of accuracy as that of results using the entropy-based approach for anatomical images registration in [122]. For example, the mean (±SD) registration error for SCR was 1.7±0.3 mm, compared to MI (2.8±4.3 mm), using intensity bins of 128 although no statistical different between SCR and MI was found. This may suggest that variance-based methods and NMI are less dependent on the quantization of the joint histogram when sparse sampling and 256 numbers of bins are used. It is worth to note that registration deteriorated drastically when MI was used.

Using the unpaired t-test, the mean registration error of MI was found to be significantly different at 5% level from all other measures using 256 bins. This indicates MI is more sensitive to the change of numbers of intensity bins than other measures and more likely leads to misregistration. To further illustrate the results of registration, summary

statistics are shown in the Table 4.2 for 128 and 256 intensity bins respectively. The median errors for MI, NMI, CR1, CR2 and SCR registration were almost identical to each other. However, the mean (±SD) errors of MI increased from 2.8±4.3 mm to 9.6±16.4 mm for the case of 128 bins and 256 bins. The increase of mean registration error for MI is probably due to major failures at 256 and 128 bins, as shown in Figure 4.4, when sparse sampling was used for registration.

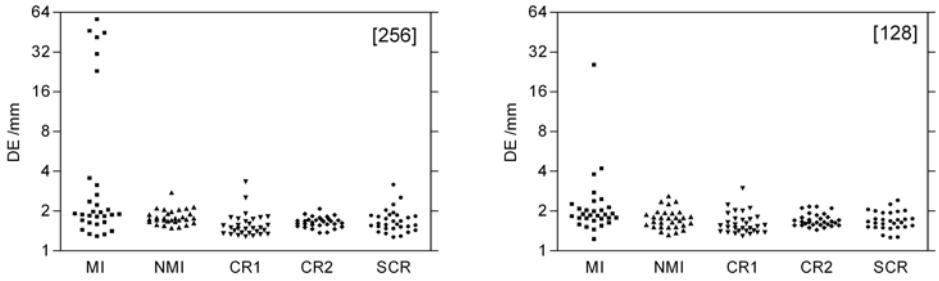


Figure 4.4: Scatter plots of registration error of T_1-T_2 MR registration for various similarity measures (MI, NMI, CR1, CR2 and SCR) using default implementation, with modified default sampling scheme: S1 (left) and S2 (right). S_1={[8,8,1|256], [4,8,1|256], [4,4,1|256]}, S_2={[8,8,1|128], [4,8,1|128], [4,4,1|128]}.

	MI		NMI		CR1		CR2		SCR	
DE in mm	S_1*	S_2	S_1	S_2	S_1	S_2	S_1	S_2	S_1	S_2
Mean ±SD	9.6±16.4	2.8±4.3	1.8±0.3	1.8 ±0.3	1.6±0.4	1.7±0.4	1.7 ±0.2	1.7±0.2	1.7±0.4	1.7±0.3
Median	1.9	1.9	1.8	1.7	1.6	1.5	1.7	1.7	1.6	1.6
Max	56.4	25.5	2.8	2.6	3.3	3.0	2.1	2.2	3.1	2.4

Table 4.2: The statistics measures of registration error using a modified default implementation: S_1={[8,8,1|256], [4,8,1|256], [4,4,1|256]}, S_2={[8,8,1|128], [4,8,1|128], [4,4,1|128]}. The asterisk denotes the presence of catastrophic outliers.

Multi-level registration

Scatter plots of the registration results for the three subsampling schemes are given in the Figure 4.5. SCR is proposed as an extension of the correlation ratio for rigid registration of multi-modality images, by incorporating symmetrical property. Surprisingly, SCR shows no failure as compared to other measures even when a large sampling factor is used for multi-modality MR image registration. The results support the assertion that variance-based measures (correlation ratio and symmetric correlation ratio) are less sensitive to subsampling while entropy-based measures (mutual information and normalized mutual information) show more pronounced local extrema (Figure 4.5). Similar observations were reported by [181]. In particular, SCR is better behaved than other similarity measures when images are sparsely sampled. This is an important consideration in practical application of registration where speed is one of the

prime factors in clinical applications. Table 4.3 shows statistics of the MR registration for the three sampling schemes and the four similarity measures (MI, NMI, CR and SCR). The robustness measures are also included.

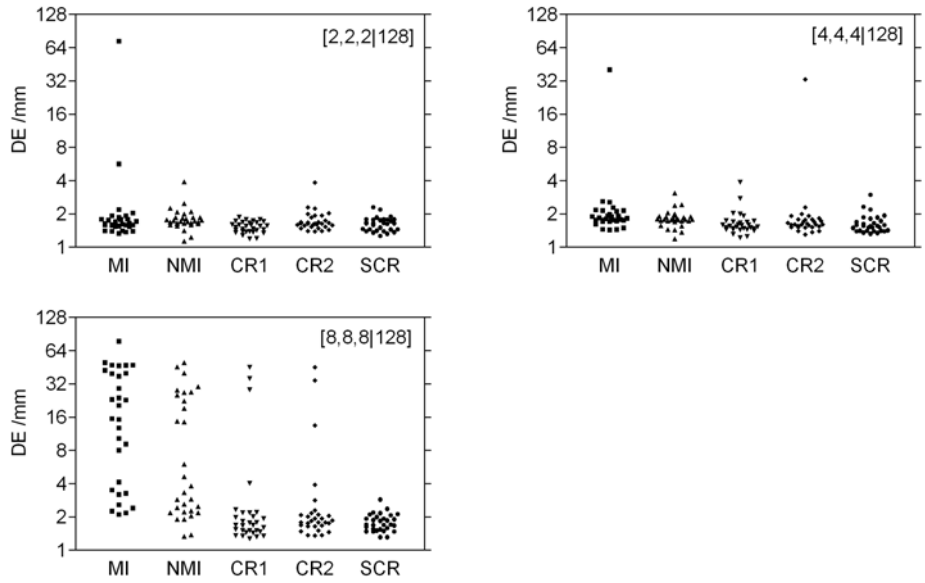


Figure 4.5: Effect of sampling scheme on MR registration carried out with different similarity measures. Individual registration errors are shown (in logarithm scale) for each similarity measure. The numbers 2, 4 and 8 denotes the various sampling scheme. Except the sampling schemes and 128 intensity bins, default implementation was used for registration. Same maximum is used in vertical scale for all graphs for comparison.

	Sampling scheme	Mean ±SD (mm)	Median (mm)	Max (mm)
MI	2	4.2±13.1	1.6	73.2
	4	3.2±7.0	1.8	40.3
	8	23.0±20.1	17.9	77.8
NMI	2	1.8±0.5	1.7	3.9
	4	1.8±0.4	1.8	3.1
	8	13.1±14.7	3.6	49.9
CR2	2	1.8±0.5	1.6	3.8
	4	2.7±5.7	1.6	33.0
	8	4.8±9.9	1.8	45.5
SCR	2	1.6±0.2	1.6	2.3
	4	1.6±0.4	1.6	3.0
	8	1.8±0.3	1.7	2.9

Table 4.3: The statistics measures of registration error as a function of sampling scheme. The numbers 2, 4, 8 represent sampling rates corresponding to [2,2,2|128], [4,4,4|128] and [8,8,8|128]. For each sampling scheme, default implementation and a maximum of 300 iterations is used.

Figure 4.5 also shows that the performance of registration depends on the choice of the similarity measures and sampling schemes. MI gives rise to large mean registration errors compared to other measures. When the registration error of MI is compared with multi-level implementation (e.g. compare [4,4,4|128] with ([8,8,1|128], [4,8,1|128],

[4,4,1|128])), the mean (±SD mm) registration error deteriorated from 2.8±4.3 to 3.2±7.0 while maximum error (mm) increased from 25.5 to 40.3. A wide spread error distribution was found for MI. The results revealed that, except in the case of a less sparse sampling in [2,2,2] and [4,4,4], a large median error was found in MI (17.9 mm) using [8,8,8]. SCR performed consistently better than MI, NMI, CR1 and CR2. SCR gave smaller mean, median and the maximum registration errors, implying that it is less sensitive to variations in the number of bins than other measures under investigation. The performance of SCR in registering anatomical and functional images in clinical studies is investigated in the following section.

4.5 Registration accuracy in clinical data with real misalignment

Further evaluations were performed to investigate the performance of SCR in the more difficult alignment of SPET to MR. In this case, experiments were performed using clinical data with actual misregistration. For the purpose of validation, external fiducial markers were attached individually to the subject's skin. The objective of the experiments was to investigate automatic SPET-MR image registration and, in particular, to compare the performance of the five similarity measures derived from the joint histogram algorithms: SCR, CR1, CR2, MI and NMI. For CR1 and CR2, variance of SPET and MR was used to normalize the conditional variance respectively (also see Subsection 3.5.3). The tests were designed to compare the accuracy and robustness of registration obtained using sparse sampling and using full sampling, as well as to examine the effectiveness of subsampling in increasing the speed of the registration without adversely influencing registration accuracy.

4.5.1 Experimental setup

Image data

For validation of the registration on clinical data, 13 subjects were scanned with both SPET and MR. Two healthy volunteers and eleven patients under medical or psychiatric investigation participated. Six fiducial markers visible in both SPET and MR scans were attached to each subject's skin. The volunteers and patients were given intravenous injections of 500 MBq and 750 MBq of ^{99m}Tc-HMPAO, respectively. The radiopharmaceutical was technetium-99m hexamethyl-propylene amine oxime (^{99m}Tc-HMPAO), widely used as a tracer for brain perfusion imaging. Each subject lay quietly

in a darkened room for at least 5 minutes after injection. The SPET scan was performed within the first hour, followed by an MR scan within 4 hours after injection. All subjects provided informed consent. The SPET and MR images datasets were provided on CD-ROM by the Queen Elizabeth Hospital, Adelaide. In each case, the MR image was the reference image while the SPET data formed the floating image.

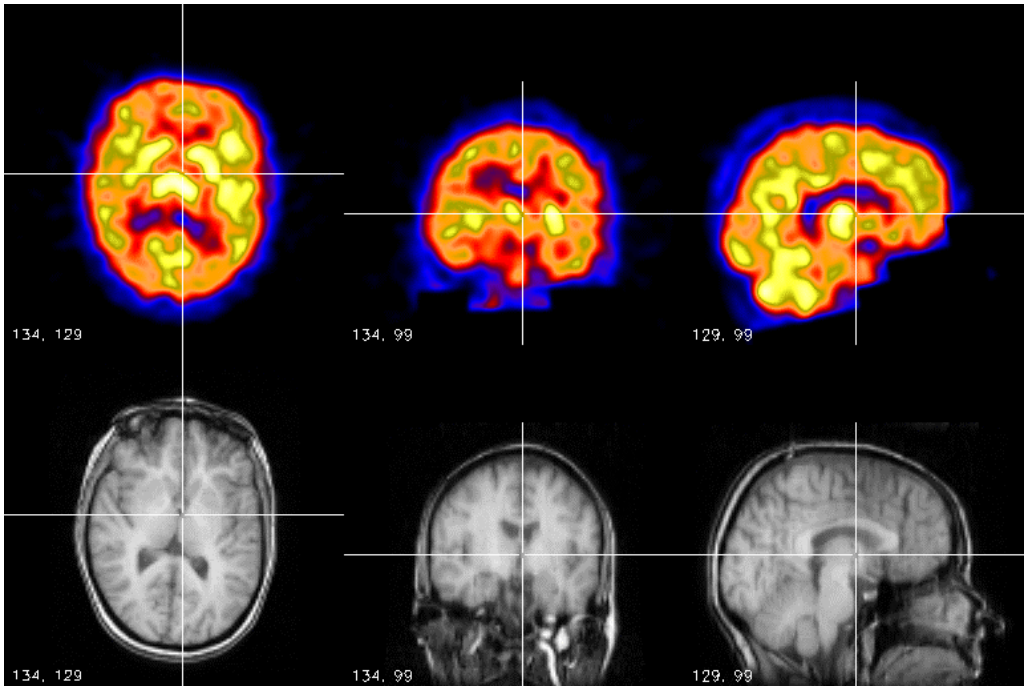


Figure 4.6: Three orthogonal slices from the original SPET (^{99m}Tc-HMPAO) dataset are shown in the top row. The point sources and the facial activity were edited out from the SPET dataset. Corresponding slices from the original MR image dataset are shown in the bottom row. The cross-wires give the relative orientations of the planar cuts. For the purpose of the display, the MR image was resampled to give cubic voxels.

Image acquisition and preprocessing

The SPET studies were performed using a triple-head gamma camera (Trionix Research Laboratories) with ultra-high resolution fan beam collimators. The system's spatial resolution is 8 mm full-width at half-maximum. The following acquisition parameters were used: 128×64 matrix and 90 or 72 stops over 360°. The projection data were passed through a 6th-order Butterworth filter with a 0.7 cycles/cm cut-off before image reconstruction. Transaxial SPET images were reconstructed using filtered back-projection with attenuation correction (attenuation coefficient 0.12/cm) to reconstruct images with 3.6 mm cubic voxels. In this particular experiment, MR images were acquired using a GE Signa 3D spoiled-gradient sequence with a 30° flip angle, inversion-recovery sequence with the inversion time T_R= 500 ms and T_E = 1.5 ms, sampled into a 256×256 matrix of 124 slices. Each slice was separated by 1.5 mm from

center-to-center. The 256×256 matrix was reduced to 128×128 giving pixel size 1.9 mm in eleven subjects and 1.72 mm in the remaining two subjects. Figure 4.6 shows sections of the SPET and MR images in three orthogonal directions.

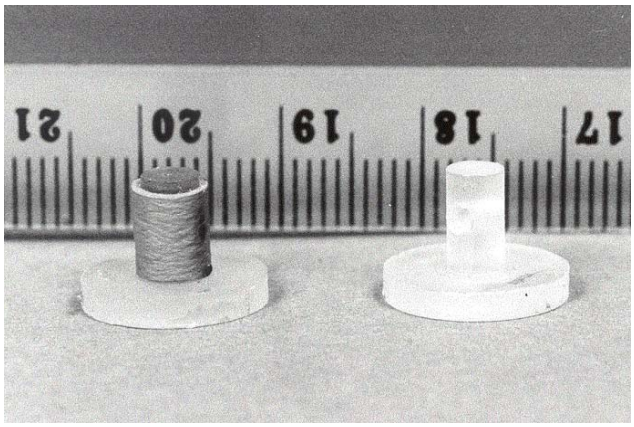


Figure 4.7: Two fiducial markers are shown against a millimeter scale. The marker on the left is fitted with a rubber sleeve. (Image courtesy of Dr. Leighton Barnden)

Fiducial markers

Each fiducial marker (as shown in Figure 4.7) was contained in a cavity (2 mm in diameter and 3.5 mm deep) within a Perspex™ cylinder attached to a Perspex circular base (12.5 mm diameter). Each cavity was loaded with 6 micro-litre of fluid containing 80 kBq ^{99m}Tc mixed with MR contrast agent (gadodiamide). The fluid was prevented from leaking by a tight-fitting rubber sleeve. Double-sided adhesive tape was attached to the base for attaching the marker onto the subject's skin. After the skin at each point had been cleaned with alcohol and allowed to dry, the markers were firmly pressed onto the skin using the adhesive. The six fiducial markers were widely spaced around the head. They were attached to each subject in pairs, 80 mm apart on the forehead, the ear and the crown. When attached, the fluid was 6 mm from the skin. Since the fiducial markers were attached to the subject's skin, SPET could be acquired without increasing the scan radius or altering spatial resolution. However, displacement of markers between scans could introduce systematic errors. In the rest of this chapter, the fiducial marker is referred to as a point source and its position is estimated by its centroid. The SPET image, from which the point source and facial activities are removed, will be referred to as the edited SPET.

Fiducial marker localization

Attenuation correction in SPET influenced the point source blur and increased errors in centroid localization. Hence, SPET without attenuation correction was used to determine the centroid location by a "center-of-mass" algorithm [278]. The MR and SPET were of different voxel sizes, matrix dimensions and coordinate systems. To allow the spatial position vector of each fiducial marker in SPET and MR to be consistently measured before and after registration, a common coordinate system was used in both of the images. The coordinate system used in the study is right-handed so that when looking from a positive axis to the origin, a positive rotation is counterclockwise. This coordinate system is the physical coordinate system of the selected image data. For example, for a transverse section of an image displayed in a graphic window, the x-axis runs across the display, the y-axis is vertical, and the positive z-axis extends out from the display to the viewer. A 90° positive rotation about the z-axis transforms the x-axis to the y-axis. Spatial coordinates are integers, ranging from (0,0) at the bottom-left corner to (127,127) at the upper-right corner of the transverse image.

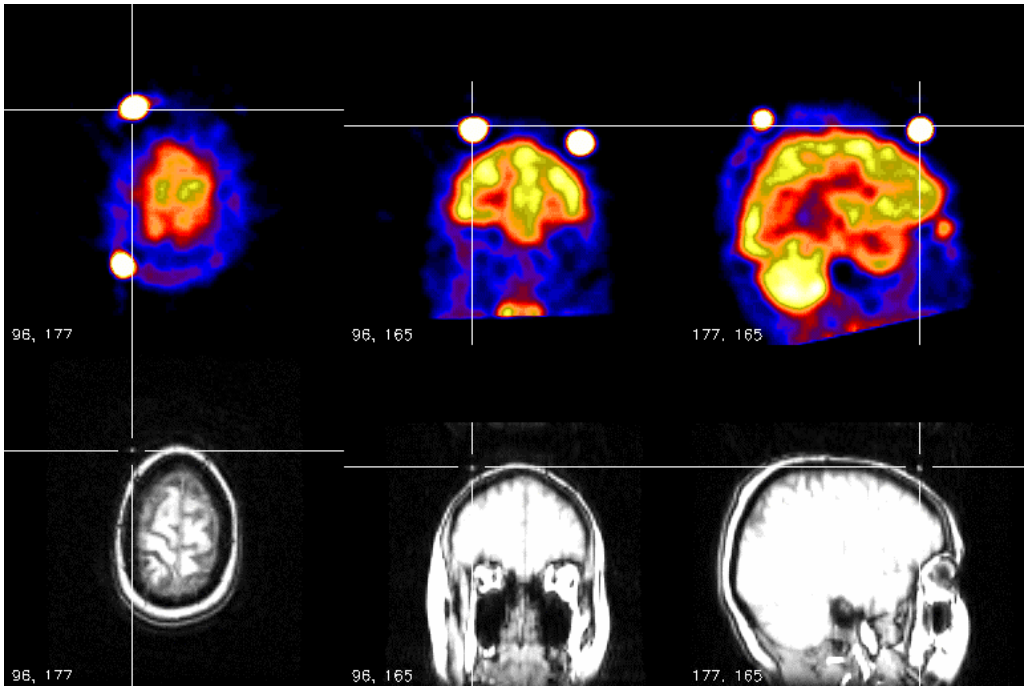


Figure 4.8: A screen shot from the GUI used to localize fiducial markers. The markers are visible in SPET (top row) and MR (bottom row). The scale has been transformed to enhance the appearance of the markers in SPET and MR. Note that the scale is over enhanced to visualize the markers in MR at the expenses of losing contrast. An operator initially, and then refined using intensities within a predefined cubic voxel estimated the centroid of a cylindrical marker. The current position of the cursor is indicated by the coordinates at the bottom left in each of the display windows.

Figure 4.8 shows the MR and SPET images with fiducial markers and the cross-wires manipulated by the user to measure the position of the markers. The center location of each fiducial marker was estimated in the images by an interactive program written in IDL™ (Interactive Data Language, Research Systems Inc, Boulder). The initial position of each marker was estimated in SPET and MR from cursors positioned by an operator in three orthogonal planes. To refine the location for each marker, a "center-of-mass" algorithm was applied to the voxels within a cube around the identified marker positions [278]. The size of the cube was 7×7×7 voxels and 3×3×3 voxels, respectively, in SPET and MR images. A total of 67 fiducial markers from 12 subjects (one data set was corrupted) were used to compute the registration error. A number of markers were dislodged between SPET and MR scans. Six markers were seen in the images of 8 subjects, two markers in 2 subjects and three markers in one subject. Once each retrospective registration was completed and T_θ was obtained, the residual displacement between the corresponding fiducial markers in the SPET and MR images could be computed easily. In order to facilitate the computation of the fiducial registration error, image preprocessing was performed to convert the dimension of each SPET to be the same as the corresponding MR before determining the spatial position of the fiducial markers.

Experimental parameters

In each pair of images, the edited SPET image was registered to the MR image. In total, 60 registrations were performed, 12 subjects for each of the five similarity measures. To evaluate the performance of each method, FRE (Subsection 4.2.2) was computed directly as a measure of the registration accuracy for each similarity measure without referring to the Procrustes approach [116] (gold standard registration). The average FRE for the 3D displacement error of all fiducial markers over all subjects was computed. The median and maximum FRE over all subjects were also calculated. The experiments were designed to examine the influence of subsampling on the accuracy and robustness of registration, and to determine whether subsampling of the data could be used to increase registration speed without adversely affecting the registration accuracy for SPET-MR registration. The default sampling/quantization scheme (Subsection 4.3.5) was varied to assess the effect of subsampling and intensity rebinning on SPET-MR registration. The following sampling strategies were used: (S1) default subsampling

strategy; (S2) [8,8,1|24], [4,4,1|32], [2,2,1|64]; (S3) [1,1,1|24], [1,1,1|24], [1,1,1|24] and (S4) [1,1,1|128], [1,1,1|28], [1,1,1|128]. For each of the 12 subjects, rigid-body registration of SPET to MR was carried out using each sampling strategy. In S1 and S2 they both started equally "coarse" but a "finer" strategy was used in S2 at the third level to determine if registration accuracy could be improved. S3 and S4 were intended to look at quantization effects.

4.6 Results

4.6.1 Comparison of registration accuracy at default setting

Registration accuracy is indicated in Figure 4.9 where a box-whiskers plot is used to illustrate the variation of mean FRE for the different similarity measures. When comparing the results for CR2 and MI, the plots both show a relatively large spread of error (i.e. large standard deviation) compared to NMI, CR1 and SCR. The results revealed that SCR, CR1 and NMI had similar performance and is confirmed by statistical test at 5% (one-way nonparametric ANOVA). The default settings gave better registration than CR2 and MI. For example, the difference of mean FRE was found to be significant for both CR2 and SCR, MI and SCR (two-tailed paired t-test, $p < 0.05$). However, MI and NMI were not found to be significantly different ($p = 0.1039$). The large maximum errors indicate that CR2 (13.1 mm) and MI (9.6 mm) are less robust than the other three methods.

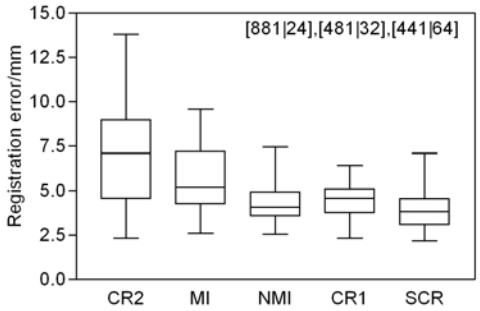


Figure 4.9: Distributions of FRE in mm are shown for SPET-MR registration against similarity measures. The line drawn across each box indicates the median of the data. The bottom and top edges of the box mark the first (25th percentile) and third (75th percentile) quartiles, respectively. The vertical line extends above and below the box to show the highest and lowest values. Default setting is used to assess the registration accuracy.

Table 4.4 further illustrates the means for FRE across the five methods based on the default setting for registration. The mean error decreases for CR2 to SCR from 7.5 mm to 3.9 mm. The mean error was supplemented by the median, maximum and standard

deviation of FRE. For entropy based methods, NMI performed better than MI over the 12 subjects. For the variance-based approaches, the mean FRE was 7.5 mm for CR2, 4.4 mm for CR1 and 3.9 mm for SCR. The fiducial displacement error in x-, y- and z-directions are illustrated in Figure 4.10.

	CR2	MI	NMI	CR1	SCR
Mean ± SD (mm)	7.5±3.3	5.6±2.3	4.4±1.4	4.4±1.1	3.9±1.3
Median (mm)	7.3	5.2	4.1	4.6	3.8
Max (mm)	13.1	9.6	7.5	6.5	7.1

Table 4.4: The table shows the SPET-MR fiducial registration errors averaged over 12 clinical data sets using the default sampling scheme: [8,8,1|32], [4,8,1|48], [4,4,1|64]. No prior smoothing of the data was applied before registration.

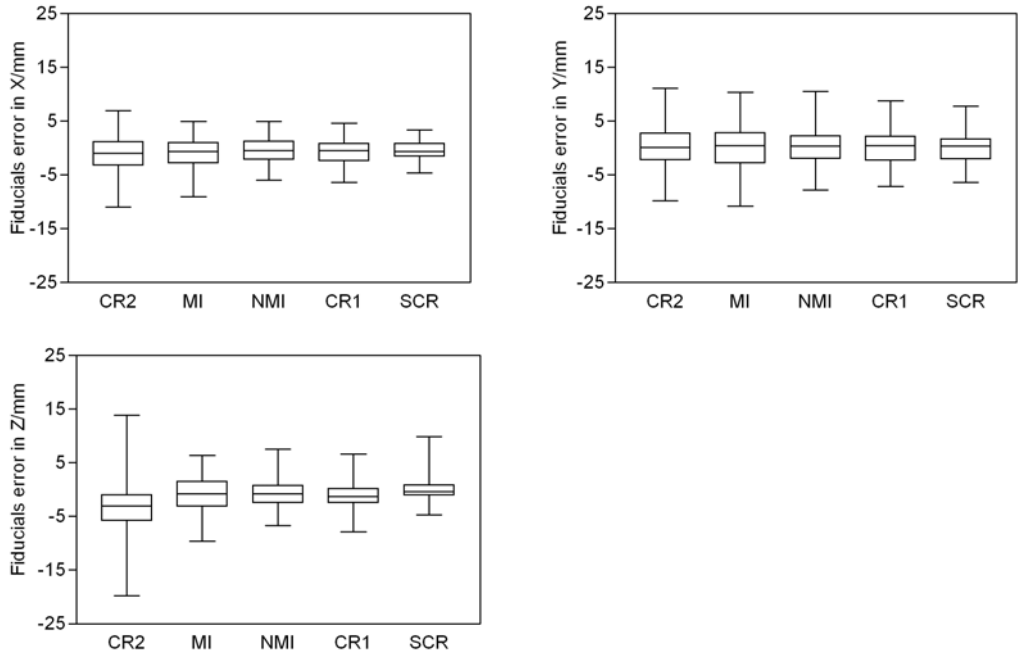


Figure 4.10: The distribution shows the variability of the x (left), y (middle) and z (right) displacement errors for SPET-MR registration across all subjects. The middle line shows the median values.

In general, most of the mean errors in each direction were small, revealing that the registration errors were mainly random in nature. Nevertheless, it was found that the largest mean and median error observed is generally in the z-direction (out-of-plane) and smallest in the y-direction (in-plane) irrespective to the similarity measures. For example, the mean and the maximum FRE of NMI in the z-direction were 0.6 mm and 0.8 mm, and the corresponding values in the y-direction were 0.1 mm and 0.4 mm respectively. These observations can be explained by the fact that the registration was less well conditioned in out-of-plane translation as the facial activity was not included in the edited SPET for registration. CR2 shows the largest displacement error while MI

is marginally worse than NMI, CR1 and SCR in the negative x- and z-direction. Based on the results of registration, the maximum rotational error was found around z-axis and the maximum translational error was along z-axis, which were 15º and 22.5 mm respectively. This may suggest that the most prominent misregistration was probably due to rotation around z-axis and translation along the z-direction.

4.6.2 Effect of sampling and quantization

The mean FRE obtained when using full sampling S4 and 128 intensity bins was taken as a reference. In the full sampling S3 and S4, simplex minimization was restarted after 100 iterations by reinitializing the registration using the parameters obtained in the last result so as to avoid a local minimum. The results of registration accuracy for the clinical data sets are given in Table 4.5, two sparse sampling schemes S1 and S2, and two full sampling schemes S3 and S4 are compared over the five similarity measures. The first observation of the results show that the mean FRE is roughly identical between similarity measures, except MI, for all sampling schemes. Using two-tailed unpaired t-tests, the difference between CR1 and similarity measures (NMI, MI, CR2 and SCR) was significant at 5%. For a given sparse subsampling scheme, CR2 and MI are found to give worst registration accuracy compared to NMI, CR1 and SCR. For example, mean FRE (±SD) in mm for CR2, MI, NMI, CR1 and SCR are 7.5±1.3, 5.6±3.3, 4.4±2.1, 4.4±1.4 and 3.9±1.1 respectively. SCR was again found to give the smallest mean FRE. The results for CR2 and MI indicated that they were least accurate with respect to the three methods regardless of the subsampling scheme used. For full sampling S4, when comparing the results of NMI to MI, SCR to CR2 and to MI, significant differences were found ($p < 0.05$). The results of the SCR and CR1, SCR and NMI did not differ significantly.

Mean (±SD) FRE (mm)	CR2	MI	NMI	CR1	SCR
S1 Sparse sampling	7.5±1.3	5.6±3.3	4.4±2.1	4.4±1.4	3.9±1.1
S2 Sparse sample	7.2±1.2	4.8±3.5	4.0±1.5	4.1±1.0	3.8±1.1
S3 Full sampling	6.8±1.1	4.7±3.2	4.0±1.5	4.0±1.0	3.8±1.0
S4 Full sampling	7.0±1.0	4.7±3.2	4.0±1.2	4.1±1.0	3.9±1.2

Table 4.5: The table shows the mean FRE error (in mm) of SPET-MR registration over 12 clinical data sets using various sampling scheme: S1 denotes default scheme; S2 denotes {[8,8,1|24], [4,4,1|32],[2,2,1|64]}; S3 denotes {[1,1,1|24], [1,1,1|24], [1,1,1|24]}; S4 denotes {[1,1,1|128], [1,1,1|128] and [1,1,1|128]}.

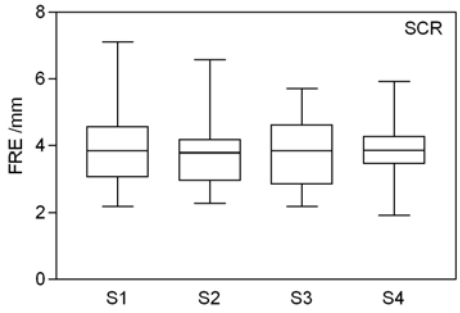


Figure 4.11: FRE (mm) for SCR using different subsample schemes. S1: default subsampling strategy; S2: {[8,8,1|24], [4,4,1|32], [2,2,1|64]}; S3: {[1,1,1|24], [1,1,1|24], [1,1,1|24]} and S4: {[1,1,1|128], [1,1,1|28], [1,1,1|128]}.

The second observation is that, based on the measurement, no significant difference at 5% level (two-tailed unpaired t-test, $p > 0.05$) was found between sparse subsampling S1 and S2 or full sampling S3 and S4 regardless of NMI, CR2, CR1 and SCR. As an illustration, typical results for SCR are given in Figure 4.11showing comparable mean FRE over the sampling schemes. Therefore, a finer subsampling strategy at sparse sampling or intensity quantization at full sampling is not critical for multi-level SPET-MR registration when NMI, CR1 and SCR were used. This results show that, except for MI, the registration accuracy obtained when sparse sampling S1 and S2 was used show little change in mean FRE when compared to the full sampling schemes S3 and S4, and this is confirmed statistically and no significant difference was found (two-tailed unpaired t-test, $p < 0.05$). This finding is important from a practical standpoint that image registration process can be accelerated using sparse sampling without adversely affecting accuracy. In this particular study a registration speed-up factor of about 22 was achieved by subsampling S1 without apparently deteriorating the registration accuracy when compared to the registration accuracy using full sampling S4.

Sampling schemes	Time (min)	Acceleration factor
S4: [1,1,1|128], [1,1,1|128], [1,1,1|128]	18.1	1.0
S3: [1,1,1|24], [1,1,1|24], [1,1,1|24]	12.3	1.5
S2: [8,8,1|24], [4,4,1|32], [2,2,1|64]	1.3	14.2
S1: default subsampling strategy	0.8	22.0

Table 4.6: Typical processing times for SPET-MR registration as measured by a stand alone PC (666 M Hz and 192 Mbytes of RAM) for the four subsampling schemes (S1, S2, S3 and S4). Registration method SCR and default setting was used for image alignment.

Table 4.6 lists the computation time for SCR registration over the four sampling schemes for the same default setting. For automatic image registration, typical registration times (in minutes) were 0.8, 1.3, 12.3 and 18.1 for S1, S2, S3 and S4

sampling schemes respectively. The results of the experiments also suggest MI is more sensitive to the sparse sampling than the other registration measures. It is also interesting to note that, for the full sampling, the number of intensity bins had little influence on the accuracy. There was no significant difference at 5% level (two-tailed paired t-test, $p = 0.05$) between mean FRE for full sampling using 24 bins and 128 bins for all the five measures. When comparing the maximum error across the five similarity measures, full sampling improved registration by reducing the maximum error irrespective to MI, NMI and SCR. For example the maximum error decreased from 7.1 mm for S1 sparse sampling to 5.9 mm for S4 full sampling in SCR. This implies the full sampling is more robust although no significant difference in mean FRE was found between sparse and full sampling for each similarity measure. The scatter of FRE was measured by the standard deviation (SD) of the error in the table. MI shows a relatively large spread of error (i.e. large standard deviation) compared to NMI, CR1 and SCR, which imply that MI was less precise for registration. This could also be attributed to the fact that MI was more likely to be affected by sparse subsampling.

4.7 Discussion

In this chapter, five joint intensity histogram based similarity measures were evaluated to study the effect of subsampling and multi-level approach on registration accuracy. Two similarity measures, mutual information and normalized mutual information, are based on entropy. The other three measures using variance-based registration are the correlation ratio (two versions depending on which image is used for normalization of the conditional variance) and the symmetrical correlation ratio.

For the variance-based similarity measures, looking at the results for T_1-T_2 MR images registration in Figure 4.5, SCR has a slight but statistically significant advantage in "robustness" over its two components. The plots are most dramatic, which shows that CR1 and CR2 both have several outliers but SCR has none. When a unilateral failure occurs, say in CR1, it implies that CR1 has a local maximum, probably far from the true maximum. Assuming that CR2 registers within acceptable limits and it has just one maximum, therefore the local max will be half the relative magnitude in SCR. This may be enough for a local optimization algorithm to avoid getting caught there, which aids in explaining why SCR produces a slightly better registration over either CR1 or CR2. In the case of dissimilar images, the results of SPET-MR registration (Figure 4.9)

show that CR2 is substantially poorer than either of CR1 or SCR. It suggests the order of image used for normalization is very important if images are substantially different. This is premised on the view of SPET activity as being generally confined to anatomical boundaries but not necessary uniformly distributed within those boundaries. For example, if the high activity areas (hot spots) of SPET are within grey matter when the SPET and MR are aligned, the regions induced in MR will fall within grey matter intensity range. Ignoring resolution effects and some intrinsic variation, it is reasonable to expect that the hot spots in SPET map to a uniform MR region. Variance in such regions in MR is low which, in turn, produces a high CR value (Eq. 3.23). It means that, for CR from SPET to MR, voxel intensity in MR is highly predictable by the voxel intensity in SPET. However, for CR from MR to SPET, if the entire grey matter volume in MR is considered, the region this induces in SPET will have highly variable intensity and therefore high variance and low CR. This implies that the voxel intensity in SPET is less likely to be predictable by the voxel intensity in MR. In SPET-MR registration, SCR comprises the performance of the CR from MR to SPET and verse versa (i.e. CR from SPET to MR), and the performance of SCR may not better than the higher of the two CRs. Thus it may not be the best approach for SPET-CT registration. On the other hand, the registration performance of CR1 looks similar to SCR. In other words, registration error would be comparable if CR1 is used (CR from SPET to MR) instead of SCR. However, the mean FRE was found to be consistently less than CR1 although their difference is not statistically significant (Table 4.5). The slight better performance of SCR than CR1 may suggest that it gives a smoother registration function and introduce less spurious local maxima compared to CR1.

The results of the experiment on subsampling and intensity rebinning suggest that SCR is less affected by sparse sampling and gives a more robust registration than other similarity measures for sub-image block registration. Similar observations were obtained in [181,182]. Interpolation induced artefacts, appearing as a pattern of local extrema in the similarity measure, are observed at full sampling [1,1] irrespective to MI, NMI, CR and SCR. Local minima were found when translation is at the integer multiples of a slice thickness due to the interpolation artifacts [219]. Here, the term interpolation artefact refers to the effect of interpolation on the similarity measures. This is different from the effect of interpolation on image quality. The results indicated that the effect of artifacts due to sparse sampling depends on the choice of the similarity

measure. In the case of sparse sampling [4,4], MI and NMI suffer more spurious local maximum than both of CR and SCR. The entropy based approach appeared to be more sensitive to the subsampling of images and this helps to explain the inferior performance of MI compared to SCR.

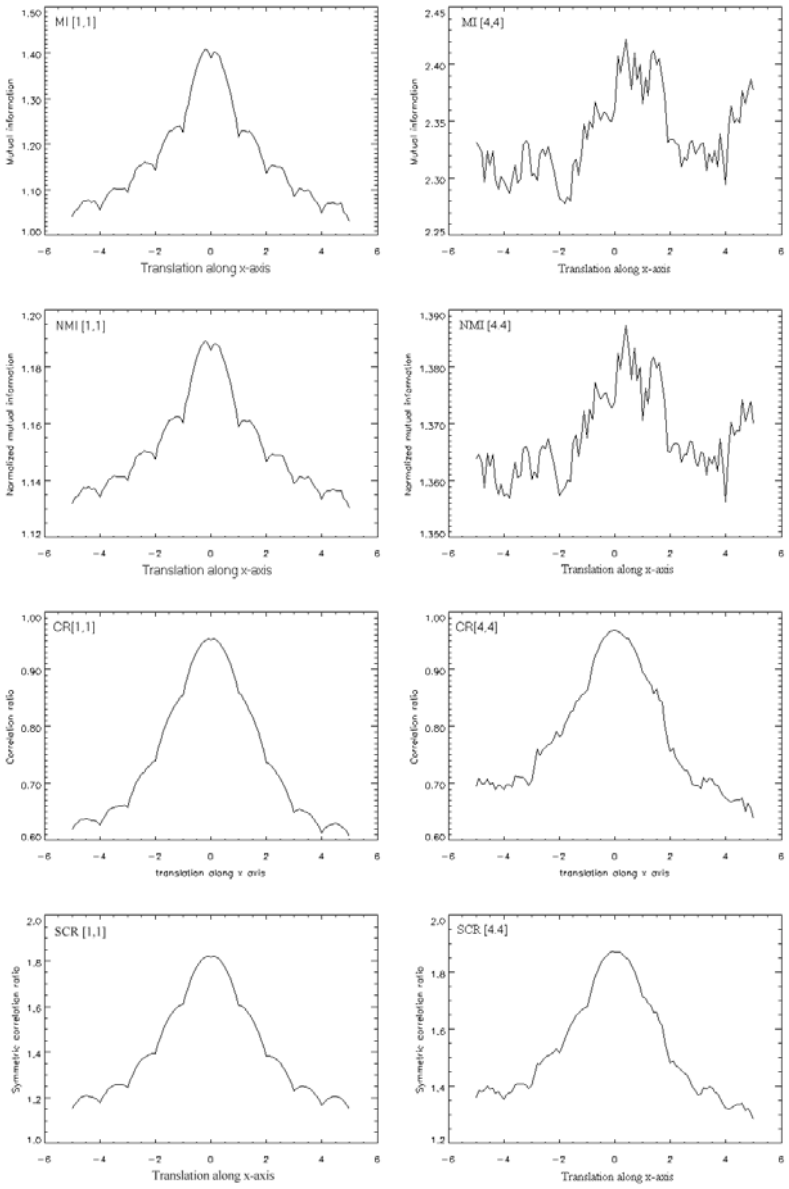


Figure 4.12: Example of interpolation effect for various similarity measures at full resolution [1,1] and sparse sampling [4,4] where data are sampled every fourth pixel in x and y directions. The measures are evaluated for a pair of 2D MR T_1 and T_2 images of size 128×128 and bins size 256. The position of zero translation is the matched position. Note the variation of the entropy based similarity measure as sampling factor increased to [4,4] when compared to [1,1]. The entropy-based methods are shown to be more sensitive to subsampling. To magnify the variation of similarity measures versus translation, the units of vertical axis in all plots are different

This observation can be understood by referring to the Figure 4.12 where variation of SCR is found to be almost independent of sampling factor for translational matching of

T_1 and T_2 MR of size 128×128 with bins size 256. They yield with translation a much smoother function, containing fewer erroneous extrema, and lead to an increase in performance of registration using sparse sampling. For CR and SCR, they both have a roughly identical result in full sampling but SCR gives a slightly more smooth variation than CR for [4,4] sparse sampling. It may help to explain SCR gave better performance than its two components shown in Figure 4.5. A registration scheme using adaptive intensity rebinning and subsampling schemes is able to preserve registration accuracy. With adaptive intensity rebinning, the sampling strategy has a small influence on the registration accuracy of the method. However, intensity quantization has little effect on registration accuracy when full sampling is used for SPET-MR registration. To further evaluate the performance of these measures based on real clinical data with actual misregistration, fiducial markers attached to the subjects' skin were used to compare these similarity measures by measuring residual displacement error after registration. The fiducial accuracy can be determined either by direct measurement (method used in this chapter) or indirectly by referring to a gold standard registration [122,123]. The gold standard registration was achieved by minimizing the distance between the corresponding fiducial markers in the two images in a least-squares sense [116].

Subject	FRE	Max FRE	FRE_g	Max FRE_g
1	3.9	7.2	4.0	7.1
2	3.1	4.0	3.8	5.5
3	7.1	10.0	4.9	8.2
4	4.6	6.0	2.8	6.1
5	4.2	8.4	3.5	7.7
6	3.0	5.1	2.5	4.8
7	4.9	8.3	3.9	7.1
8	/	/	5.0	6.3
9	4.6	7.7	5.1	8.2
10	3.7	6.4	3.5	6.1
11	3.8	6.7	4.0	6.0
12	2.2	4.2	4.4	5.5
13	2.2	2.9	2.2	4.2
Mean±SD (mm)	3.9±1.3	6.4±2.1	3.8±0.9	6.4±1.2

Table 4.7: The table shows fiducial registration error for each subject and means fiducial registration errors over all subjects in mm for SPECT-MR registration using SCR and default implementation. Notes that the subject 8 data was corrupted in the current study. FRE_g obtained with referencing to the gold transformation is also given for comparison.

To distinguish these two cases a subscript g is added to denote FRE obtained *indirectly*. Results in Table 4.7 shows that differences of mean FRE between direct and indirect measurements were not significant statistically (two-tailed paired t-test, $p<0.05$).

Similar observation was obtained for maximum FRE obtained directly or indirectly. This implies that registration accuracy can be determined directly without necessary referring to a gold standard registration, such as Procrustes approach [116]. These results suggest that a direct measurement of spatial distance for the corresponding fiducial markers in SPET and MR after registration can be an alternative and efficient approach to assess registration accuracy. This removes the need for the extra step in establishing registration accuracy relative to a method based on the gold standard registration. The errors (FRE = 3.9 mm) obtained for SPET-MR registration can be compared with results of Collignon (3.6 mm) and Hill (3.2 mm) for PET to MR registration errors as published by West *et al.* [122]. Intrinsic registration error based on internal anatomical features provides a more realistic error measures. The intrinsic SPET to MR registration error (IRE = 2.5 mm) determined from anatomical locations using SCR in [123] is also comparable to [122], where the intrinsic registration errors for the PET-MR registration using MI was 2.4 mm. Given that the SPET resolution is worse than PET resolution, these results are quite satisfactory. The uncertainty of localization of the center of a point source distribution gives the so-called fiducial location error (FLE), which may be due to skin markers movement between scans, non-uniformity in MR scans or subject's motion during image acquisition. The FLE was monitored by the systematic error due to inter-marker transformation between MR and SPET scans. The distance from one marker to another in MR was compared to corresponding distances in SPET [123]. The distance was measured over all marker pairs to give the mean discrepancies for each subject. The discrepancy in distance gives a measure of symmetric error that is independent of the registration process. Across all subjects, most of the discrepancy was smaller than 1 mm except for subjects 3 and 7 (Figure 4.13). These results suggest that displacement of one or more fiducial markers between MR and SPET scans was small. Thus the effect of FLE in the current experiment was minimal and had little effect on the FRE measurement. Because of the discrete nature of images, the transformed image grid rarely coincides with the reference image grid and interpolation is required. Another complication is that the transformed floating image may fall outside the grid of the reference image. This problem is solved either by padding out the reference image with an arbitrary intensity, such as zero, and later discarding these data or by restricting the computation to voxels in the overlapping region [192]. The reconstruction accuracy of voxel intensity values is dependent on the interpolation kernel used (e.g. nearest neighbour [225], trilinear [123]

or partial volume interpolation [275]). It is unclear how the subsampling factors would influence the registration due to the effect of aliasing. The decrease in the number of voxels due to subsampling may yield a less smooth registration function from the joint histogram as less sample data and information from images is available [230]. Thus, a reduced number of bins are desirable so that there is at least one entry per bin when subsampling is applied to the images. However, too few bins can result in a loss of image detail. The strategy to increase the number of bins adaptively with finer sampling seems applicable for T_1-T_2 MR registration and MR-SPET registration without worsening the registration accuracy.

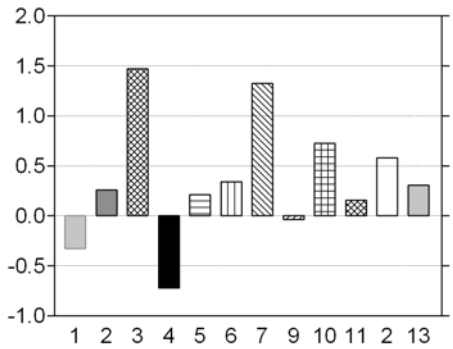


Figure 4.13: The vertical axis denotes the mean discrepancy in mm for each subject. The data set 8 is missing because the data were corrupted.

4.8 Conclusion

In summary, a symmetrical correlation ratio for 3D multi-modality registration is validated. The performance of the proposed algorithm is compared to MI, NMI and CR for T_1-T_2 MR registration and MR-SPET registration. Comparative studies of adaptive intensity rebinning and subsampling schemes over MI, NMI and SCR have been undertaken. Results suggest that the proposed method based on multi-level strategy and SCR is effective to accelerate registration without sacrificing registration accuracy. The algorithm is robust, fast and reliable, even in the case of inter-modality registration. Using SCR, the mean displacement error for T_1-T_2 MR registration and the mean fiducial registration error for SPET-MR registration are both in the order of one voxel, which is sufficient for most medical applications.

Chapter 5

SPET acquisition and a torso phantom

5.1 Introduction

Understanding characteristics of data acquisition for each modality involved in a registration problem is important to the development of registration methods and their validation. In particular, this understanding is crucial to simulation of SPET. This chapter presents background material relating to SPET acquisition. Although not as popular as neurological registration, the subject of non-brain registration is receiving increasing attention [79,81,130,160,197,279,80,280,281,282,283]. Somer *et al*. [75] studied the performance of PET-MRI rigid-body registration in soft tissue based on a point-based external marker technique and compared the results with those obtained using mutual information. Yu *et al.* [77,78] proposed a surfaces-based algorithm to register the CT and PET studies in the thorax region. Noz *et al*. [196] studied registration of CT/MRI and SPET data using a polynomial warping technique. A number of multimodality registration studies in the non-brain area were based on phantoms [254,261,260]. Based on a thorax phantom, Meyer *et al.* [186] used affine mapping and rigid-body mapping to study PET-CT and SPET-CT registration performance, respectively. Livieratos *et al.* [253] used digital phantoms to assess PET-PET registration in the abdomen. In order to lead into the discussion of a digital torso SPET phantom, this chapter sets out the principles of acquisition of SPET, describes image degradation factors that need to be incorporated in a SPET phantom, and briefly presents principal reconstruction methods. This is followed by the description of the SPET torso phantom developed for the purpose of assessing the feasibility and accuracy of SPET-CT retrospective registration in the thoraco-abdominal region.

5.2 Data acquisition in nuclear medicine

In nuclear medicine procedures, patients are administered radioactive substances for the purpose of diagnosis or therapy. In this chapter, the focus is on diagnostic imaging, with emphasis on the Single Photon Emission Tomography (SPET). The main objective of clinical SPET is to study the functioning of organs and tissues. This is achieved by administering to the patient compounds tagged with a radionuclide

(radiopharmaceuticals) whose distribution in the patient's tissues reflects their function. Radionuclide decay may be accompanied by the emission of α, β, and γ radiation. Only γ emission is useful for diagnostic imaging because α and β particles have a short range and become absorbed before they can be detected. Often, the radipharmaceutical will concentrate in a particular organ of the body, allowing detailed imaging and functional analysis of that organ. The distribution can be imaged over time to see how it changes (dynamic) or just imaged once to demonstrate its spatial distribution (static) as shown in Figure 5.1, where a planar projection view of a whole-body scan is illustrated. The essential idea is that the radioactive substance should act as a tracer for a particular physiological process. A list of the more common radionuclides used in diagnostic imaging is given in table 5.1.

Radionuclide	Photon energies (keV)	Half-life (hours)
^{99m}Tc	140	6.0
^{111m}In	172 and 247	67.2
^{67}Ga	93, 185 and 300	78.24
^{201}Tl	135 and 167	72.96
^{123}I	159 and 529	13.3

Table 5.1: Radionuclides commonly used in diagnostic nuclear medicine.

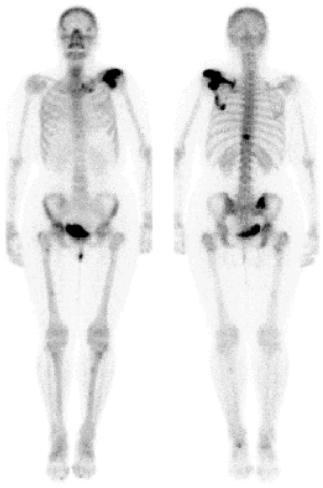


Figure 5.1: A whole body scan shows the accumulation of the tracer in bone: front view (left) and rear view (right). Metabolic bone abnormalities tend to lead to an accumulation of the radiopharmaceutical. The hot areas (dark spots) indicate possible locations of bone disease.

The radionuclide must have certain properties. It should emit only γ radiation of an energy suitable for easy detection by a gamma detector (camera) and have a suitably short half-life to reduce the radioactive exposure to the patients. In SPET imaging, the emission energy of the radiopharmaceutical lies in the general range of 100 to 400 keV.

The principal imaging device used in nuclear medicine is the scintillation gamma camera originated from the principles developed by Anger in 1958 [31]. A schematic diagram is shown in the Figure 5.2. The basic SPET system consists of one or more gamma cameras that rotate around the patient in order to acquire photons from different directions. The essential components of a gamma camera are: (1) the collimator to define the incident photon direction; (2) the scintillating crystal to transform the incoming γ photon into visible light flashes; (3) an array of photomultiplier tubes (PMT) to convert the flashes into amplified electrical pulse; and (4) associated circuitry used to determine the location and energy of detected photons [7,284,285].

The detection process of a gamma camera can be described as follows. When a γ photon is absorbed in the scintillator (usually a sodium iodide crystal doped with thallium), the scintillator atoms are raised to an excited state. They subsequently return to the ground state by the emission of visible light. These minute flashes of light provide a means of detecting γ photons. In the gamma camera, the scintillator is a large sodium iodide crystal [286,287], activated with about 0.5% of thallium oxide, abbreviated to NaI (Tl), usually circular or rectangular in shape, up to 500 mm across, and from 6 mm to 25 mm thick. Since the crystal is highly transparent to its own light emission, the visible light will escape from the crystal. As the light flashes are too small to be detected directly, they need amplification using an array of photomultipier tubes (PMT), positioned immediately behind the scintillator. The size of the output electrical pulse is directly proportional to the incoming photon energy. The most intense signal is obtained from the PMT nearest the event, and progressively weaker signals found as the distance increases. By examining the relative strength of signals from all PMTs, the position-decoding circuit determines the location of the scintillation event which, in turn, is in line with the source of the γ photon. This provides a means of identifying where in the emitting organ the original γ photon originated. The pulse height analyzer (PHA) is used to ensure the γ ray has energy within the range expected for the radionuclide used. This reduces the possibility of detecting stray radiation but is mainly intended to minimize the detection of photons that have undergone Compton scatter in the body (see Subsection 5.4.3 on scatter). The energy of the incident γ photon deposited in the crystal is estimated by summing all the PMT pulses and then feeding the sum to a PHA. If the γ

photon energy falls within a selected range of energy (referred to as the energy window), the γ photon is accepted.

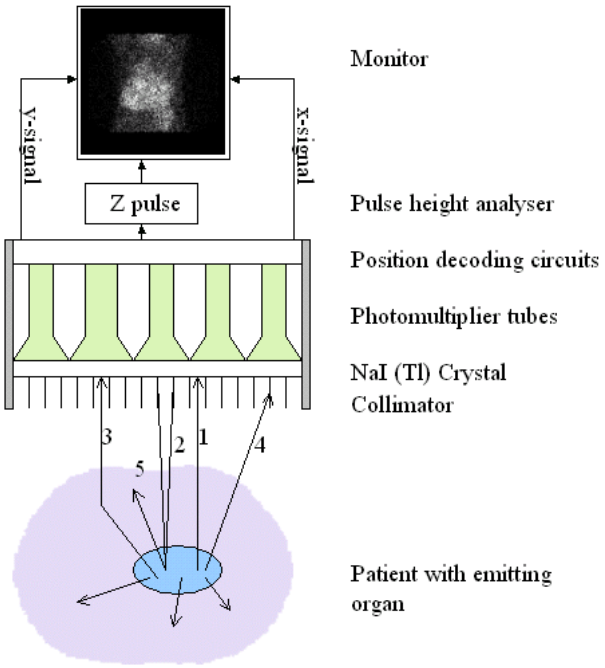


Figure 5.2: Schematic diagram of a typical Anger gamma camera, with a parallel-hole collimator is illustrated. A single planar view or "projection" is shown. A few rays are shown to illustrate typical modes of propagation. Ray 1 travels in the direction perpendicular to the collimator and is detected the crystal. Rays 2 show the range of incidence angles that will still result in detection, but the finite dimensions of the collimator hole will cause image blurring. Ray 3 is scattered towards the camera but is detected as non-scattered ray. Ray 4 is absorbed by the septa because its angle of incident is too large to pass through the collimator. Ray 5 is absorbed by the body or is scattered away from the camera.

In digital acquisition, the detector is viewed as a matrix consisting a grid of (e.g. 64×64, 128×128) square detector elements, which form the 2D digital planar image. In general, the field of view of a typical gamma camera is about 0.5×0.5 m^2. The number of scintillations or counts obtained in each detector elements over a period of time is recorded and stored in the memory of a computer. During image acquisition, the camera is placed adjacent to the patient and both patient and the camera are kept still while counts accumulate. This results in a single planar view or "projection" of the radionuclide distribution within the patient (Figure 5.2). Projection data conceal information regarding the depth at which disintegrations occur. Thus radioactive sources that overlay each other along the line of sight to the camera cannot be distinguished. It is, therefore, necessary to acquire multiple projections of the patient at different angles so that sufficient data are obtained to reconstruct the 3D activity distribution within the body. Since these scintillation crystals are incapable of defining

the incident γ photon direction, gamma cameras are provided with collimators. A collimator consists of a honeycomb with tens of thousands of holes, parallel to the axis of the collimator. The lead separations between the holes are called septa and these ideally limit detection to only those photons that travel perpendicular to the detector, thus clearly defining their incident direction. All non-axial γ photons are absorbed in the lead septa. The performance of collimator depends on the size of the collimator holes, the thickness and the length of the septa, and the source camera distance [284]. The hole size and the septa length determine the acceptance angle of the collimator. The number of detected photons per activity unit defines the sensitivity. Long septa limit the acceptance angle of collimator holes, so few γ rays are accepted (low sensitivity) but those accepted will be aligned close to the ideal projection line (a good spatial resolution). The collimator design is a compromise between spatial resolution and sensitivity [288,289]. The thickness of the septa is chosen to match the relevant energy of the γ ray being just thick enough to stop non-axial rays penetrating, but not too thick to reduce sensitivity.

5.3 Image reconstruction in SPET

In SPET, the camera rotates around the patient, with its plane parallel to the patient's long axis, acquiring 2D projections of the 3D radionuclide distribution at different angles. This section is concerned with the process of reconstructing the SPET image from these projections. In general, the camera(s) rotate about the patient along an arc spanning at least 180°, acquiring projections at a number of positions, to provide sufficient information for reconstruction. Projection images are often acquired with a matrix size of 64×64 or 128×128 picture elements. Each row in the matrix represents a projection of a single slice. A $n \times n$ acquisition matrix contains the projections of n slices at a particular angle. Therefore, the complete set of projections provides information required to reconstruct n tomographic slices acquired by the gamma camera field of view. For a typical gamma camera diameter of 400 mm, the slice thickness is approximately 6.3 mm for a 64×64 matrix acquisition. For the sake of simplicity, data acquisition of a single slice is shown in Figure 5.3. It is assumed that the intensity of radioactive decays is proportional to radioactivity distribution $f(x,y)$ and that single photon emissions are released in an isotropic manner. The slice profile $g(s,\theta)$ is defined as the total counts detected in the detector pixel located at s from the origin when the

detector is at an angular position θ. The projection $g(s,\theta)$ of $f(x,y)$ is known as the Radon transform which represents the line integral (ray sum) of the values of $f(x,y)$ along the ray inclined at an angle θ with respect to the x-axis at a distance s from the origin. The ray is defined by the line equation $u = x\cos\theta + y\sin\theta$. The Radon transform of $f(x,y)$ along the ray is given by Equation (5.1)

$$g(s,\theta) = \int_{-\infty}^{\infty} f(x,y)\,du$$
$$= \int_{-\infty}^{\infty} f(u\cos\theta - s\sin\theta, u\sin\theta + s\cos\theta)\,du \qquad (5.1)$$

The integration path variable u is measured along the ray (Figure 5.3).

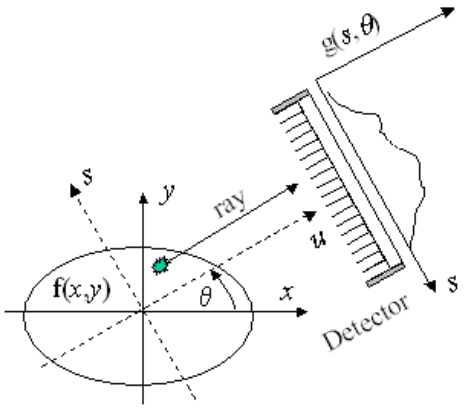


Figure 5.3: A schematic diagram of the projection geometry corresponding to a SPET acquisition with a rotation camera is shown. At each angle θ, slice profile $g(s,\theta)$ refers to the variation of intensity through (at right angles to) the slice. It gives a projection of radioactivity distribution within a transverse slice through the patient onto the detector. As the camera rotates about the patient, it acquires a 2D projection image at each θ, and each row of the projection image corresponds to a projection of the slice. A profile of a single slice is shown. The schematic diagram shows the ray sum $g(s,\theta)$ of a ray passing through the point (x,y) is identified by the angle θ it forms with the x axis, and its perpendicular distance s from the origin.

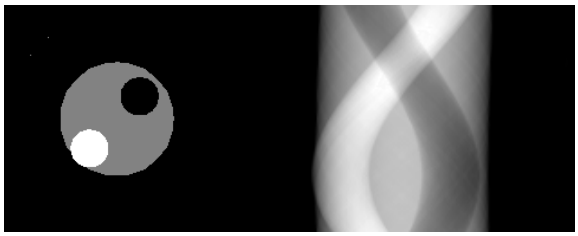


Figure 5.4: A phantom slice (left) and its corresponding sinogram (right). Each row of the sinogram is a projection of the slice at a specific view angle. The projection in the bottom row is obtained at 0^o and that in the top row is obtained at 180^o. There are 128 views (projections) equally spaced over the 180^o range. This simulation is noise free.

In discrete form, the integral represents the sum of the values f(x,y) along the ray, and is known as the ray sum. The radioactivity distribution of the slice f(x,y) is estimated from

the projection data g(s,θ). The projection data can be represented by a 2D diagram known as a sinogram. In the sinogram (Figure 5.4), the horizontal axis denotes the location on the detector and the vertical axis denotes the angular position of the detector. The sinogram gives the number of counts at each point s of the detector for each projection angle θ. The problem is: given a sinogram g(s,θ), what is the radioactivity distribution f(x,y) in the slice? Many algorithms for SPET reconstruction exist [6,32,290,291 This section briefly discuses two widely used approaches, the filtered back projection (FBP) and maximum likelihood expectation maximization (MLEM) algorithms. The mathematics of image reconstruction is treated by [292,293,294,295,296,297] and a more in-depth treatment is in [298].

5.3.1 Filtered Back Projection

It is possible to estimate f(x,y) from the measured projections g(s,θ) by a simple back projection operation defined by (5.2a). Since the projections acquired at angles between π and 2π do not provide new information in an ideal condition (e.g. without attenuation), the back projection can be limited to π radians. In the back projection process, the measured signal corresponding to each ray sum is assumed to distribute equally over all points along the ray. In other words, for a given projection angle θ, the projection g(s,θ) at location s that is corresponding to the ray sum passing through the point (x,y) in the slice is added. The process is repeated for all projection angles θ, and followed by dividing the total with the number of projections. Algorithmically [298], an estimate of the image $\hat{\mathrm{f}}(x,y)$ can be expressed as (5.2b)

$$\mathrm{f}(x,y) = \int_0^{\pi} g(s,\theta)\mathrm{d}\theta \qquad (5.2a)$$

$$\hat{\mathrm{f}}(x,y) = \sum_{k=1}^{p} \mathrm{g}(y\cos\theta_k - x\sin\theta_k, \theta_k)\Delta\theta\,, \qquad (5.2b)$$

where $\Delta\theta = \frac{\pi}{p}$ is the angular step between successive projections, p is the total number of projections acquired over π radians and θ_k is the k^{th} angular position of the detector. Figure 5.5 shows a simulation of the back projection operation.

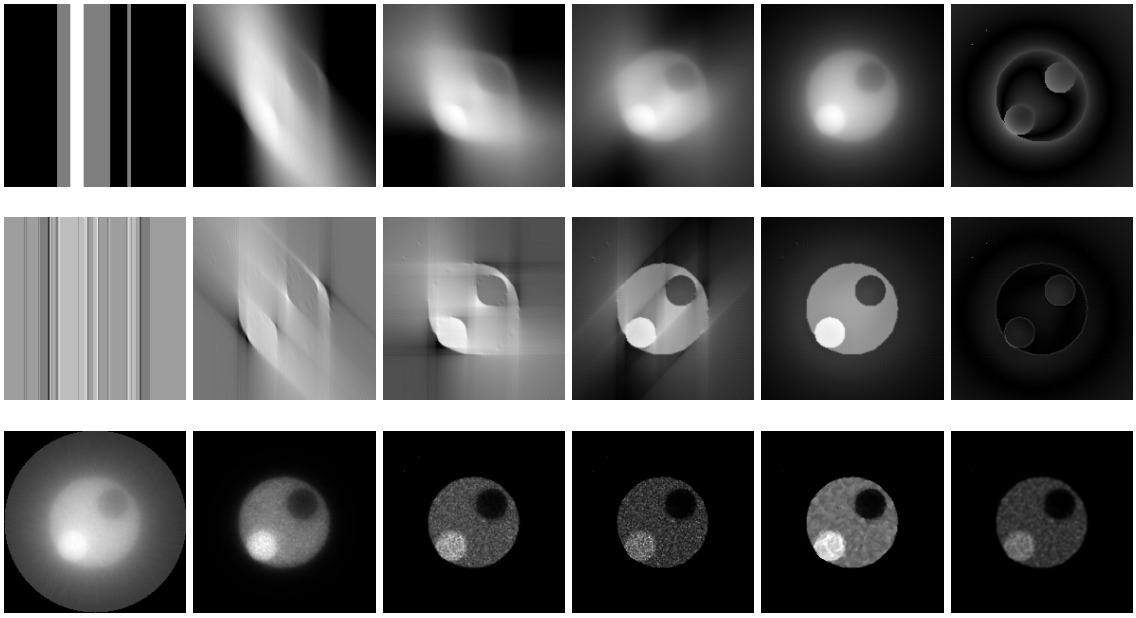


Figure 5.5: *The top row*: Direct back projection (BP) reconstruction of the phantom shown in Figure 5.4. Here an ideal case is considered, without attenuation, scattering and noise. Partial reconstructions after 1, 32, 64, 94 and 128 projections are shown (left to right). The BP reconstruction image of 128 projections (second image from the right) is similar to a blurred version of the original image. The last diagram shows the difference between BP reconstructed image and the original image.

The middle row: As in the top row, but each projection data set was pre-processed by a ramp+Hamming filter prior to back projection.

The bottom row: MLEM reconstruction of the same phantom. Poisson noise was added to projection data (128 views over 180 angles). The image noise increases as the number of iterations is increased from 1, 5, 20 and 40 iterations (left to right). The MLEM reconstruction in the second rightmost image used 40 iterations with the median root prior [299]. The rightmost image used 40 iterations with post filtering. Each image display is scaled independently.

It is interesting to note that, after back projection, the reconstructed image is not quite the same as the original image. Rather, it resembles a blurred version of the original image (Figure 5.5 top row, second image from the right) and is characterized by a loss of resolution and contrast. It can be shown that the point-spread function that describes the blurring is a "1/r" function, where r is the distance from the point source [295,297,300,301]. Consider a point source and its projections. Its reconstruction is a star like object because counts are smeared uniformly at all pixels along each non-zero ray when it is back projected. For infinitely many projections, intensity will drop off at a rate inversely proportional to distance r. As any radioactivity distribution can be decomposed into point sources, the resulting reconstruction is a convolution of the reconstructed object with the function 1/r. To reconstruct the true image from the back projected image, the "1/r" blurring factor must be eliminated. In filtered back projection (FBP), this is done by applying a filter with ramp-like frequency response (Fig. 5.6) to

the projection profiles prior to the back projection operation. Note that the bandwidth of the ramp filter is limited to a cut-off frequency. The ramp filter has the desirable effect of reducing the 1/*r* blurring.

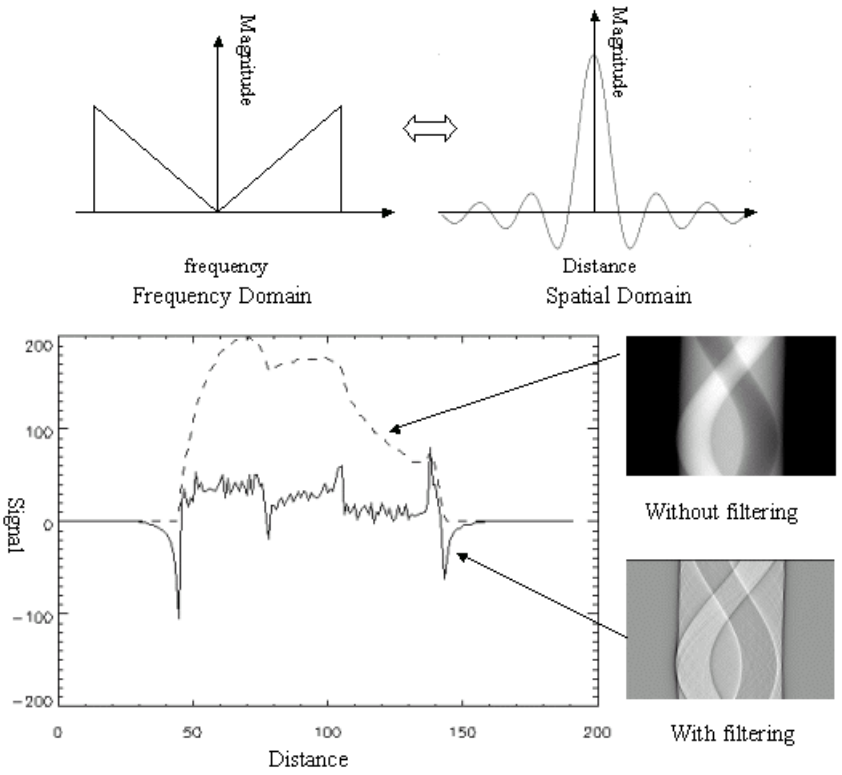


Figure 5.6: Ramp filter representation in the frequency domain (top left) and spatial domain (top right). Note the cut-off is sharp in the frequency domain and the negative side-lobe of the ramp filter in the spatial domain. Bottom: the sinograms of the slice phantom obtained with and without ramp+Hamming filtering are shown. A typical row profile extracted from these sinograms is also shown. Note the negative values of the row profile after filtering.

In FBP, each row of the sinogram is convolved with a decaying oscillating function (the impulse response of the ramp filter), which introduces negative values before backprojection. The filtered sinogram (Figure 5.6) shows two dark bands encapsulating each sine wave. A typical line profile with and without filtering is also shown. The blur is reduced because negative and positive values cancel each other near the edges of the image (Figure 5.5 middle row). In an ideal situation without attenuation, scatter, depth dependent resolution and noise, the original object f(x,y,z) is recreated faithfully. In practice, noise exists in the projection data and its frequency spectrum extends to high frequencies. Therefore, one major disadvantage of the ramp filter is that it will amplify high frequency image noise. This problem can be reduced by windowing the frequency response of the ramp filter (i.e. augmenting the high-pass ramp filter with a low-pass filter), thus providing a gentler high frequency roll-off. The rounding of the sharp

transition at the cut-off frequency will also prevent ringing artefacts in the reconstructed image [302]. However, there is a trade-off involved: by limiting high frequency component, not only the image noise is reduced but spatial resolution and image contrast may also be lost. A more detailed discussion on filters can be found in [293,295,298,301,303].

5.3.2 Maximum Likelihood Expectation Maximization

FBP is an efficient method of reconstruction, but it is based on a simple assumption that the data are simple projection ray sums. When image degrading factors such as attenuation, depth dependent resolution, scatter and Poisson noise are present, FBP approaches are likely to produce artefacts [38]. Stochastic reconstruction methods improve the estimate of the radioactivity distribution f(x,y,z) by incorporating a realistic model of the acquisition process that can account for noise, collimator geometry, non-uniform attenuation and scatter. In addition, a Bayesian formulation of the stochastic model permits the incorporation of a priori information about the radioactivity distribution [299,304]. A general discussion on iterative reconstruction algorithms in nuclear medicine can be found in [296].

The detection process in SPET can be modeled in discrete form as follows. Let the image be f(x,y,z) denoted as an image of total D voxels such that $\mathrm{f} = \{\mathrm{f}_j : j = 1,2,...,D\}$ and the voxels in f are indexed by j. Suppose there are N_θ projection angles and N bins per projection (i.e. N is the size of acquisition matrix), the acquisition data can be represented by a one dimensional vector of measurements $\mathrm{g}_i = \{\mathrm{g}_i : i = 1,2,...,m, m = N_\theta \times N\}$, where m is the number of projection angles multiplied by the number of projection bins [294]. Thus g_i denotes the number of counts in the i^{th} measurement defined over the produce space $N_\theta \times N$ and is given by a weighted sum of D voxel values in the image f, $\mathrm{g}_i = \sum_{j=1}^{D} a_{ij}\mathrm{f}_j$. The weight a_{ij} represents the contributions of voxel j to the i^{th} detector count. Stochastically, a_{ij} measures the probability of a photon emitted from voxel j being detected in i^{th} measurement in the detector. This equation can be put in a more compact form as $\mathrm{g} = \boldsymbol{A}\mathrm{f}$, where the matrix $\boldsymbol{A}$ is known as the transition matrix [296,98]. The advantage

of stochastic reconstruction is that non-uniform attenuation, scatter, depth dependent resolution and acquisition geometry can be incorporated into the transition matrix $\boldsymbol{A}$, leading to improved reconstruction accuracy [305]. The most commonly used stochastic reconstruction method is the iterative maximum likelihood expectation maximum (MLEM) algorithm proposed by Shepp and Vardi [292], which incorporated the Poisson nature of the acquired data. Their reconstruction algorithm attempts to estimate the source distribution that most likely produces the emission projection data with the highest probability. The MLEM algorithm works as follows. Initially, a uniform count in each slice voxel is assumed. Next, each voxel estimate is updated according to two steps: the E step (forward step) estimates the expected projection based on the current estimate of activity distribution; the M step (back projection) compares the expected projection with the measured projection. A discrepancy found in the M step is used to update the estimate of the activity distribution. This leads to the iterative scheme [306,307,308],

$$\hat{\mathrm{f}}_j^{k+1} = \frac{\hat{\mathrm{f}}_j^{k}}{\sum_{i=1}^{m} a_{ij}} \sum_{i=1}^{m} \frac{\mathrm{g}_i}{\sum_{j=1}^{D} a_{ij} \hat{\mathrm{f}}_j^{k}} a_{ij} \ , \qquad (5.3)$$

where $\hat{\mathrm{f}}^k$ is the k^{th} estimate of the activity distribution in the image. The MLEM algorithm converges slowly and the reconstruction becomes noisier with increasing iterations (as demonstrated in the first four images from left to right in the bottom row of Figure 5.5). Possible ways to reduce image noise are: (1) stopping the iteration process early [309]; (2) using a prior to regularize the reconstruction [299,304,307] (second last image in bottom row of Figure 5.5); (3) post filtering the reconstructed image [310] (last image in the bottom row of Figure 5.5).

One major drawback of MLEM is its long computation time when compared to the FBP reconstruction, especially when a complex projection model is involved. Hudson and Larkin [98] split the measured data set into ordered subsets (OS) and used only one subset for each MLEM sub-iteration. For example, if there were 64 projections, one might divide the projections into 16 subsets (OS16), each consisting of 4 projections. In the ordered subsets expectation maximization (OSEM), projection data are grouped into subsets of projections, ordered so as to provide maximum new information at each sub-iteration. In the sub-iteration, the same E and M steps are applied but only a subset of

the complete set of projections is used, giving a significant reduction in the processing time. One complete iteration of OSEM is defined as one pass through all the sub-iterations. Theoretically, OSEM will speed up the reconstruction by a factor which equals the number of ordered subsets. For example, a 16-subset OSEM will accelerate the reconstruction by a factor of 16 relative to the standard MLEM. OSEM has become the most popular choice of iterative reconstruction algorithm in nuclear medicine. A detailed discussion of OSEM, from a clinical perspective, can be found in [38].

5.4 Image degradation

SPET image quality degrades due to several factors [284]. These factors are caused by the finite spatial resolution of the camera (partial volume effects) and the depth dependence of resolution, non-uniform attenuation and scatter of photons in the patient, patient movement, and noise in the projection data. An overview of image degradation is given below.

5.4.1 Distance dependent resolution

The spatial resolution of a gamma camera is limited by detector properties, but is dominated by the geometric resolution of the collimator. Due to the finite collimator hole size, rays not exactly parallel to the collimator axis may still be detected by the camera (ray 2, Figure 5.2). Because the angle of incidence determines whether a ray will be detected, the camera resolution depends on the source to camera distance. The spatial resolution of a camera is usually described by the point-spread function (PSF), which can be approximated by a Gaussian curve. The spatial resolution of a camera is usually specified by the full width at half maximum (FWHM) of the PSF. The FWHM of the collimator PSF increases approximately linearly with source distance from the camera as illustrated in Figure 5.7. This effect limits the ability of the physicians to discriminate between closely located activity sources, leading to difficulty in interpreting the pathology and functional information present in the image. The spatial resolution of a camera can be improved by reducing the collimator hole diameter, but this leads to a reduction in sensitivity because of the smaller number of number of photons being detected. The effect of collimator blurring and methods of reducing the blur have been widely discussed (e.g. Rosenthal [284]). Stochastic reconstruction algorithms that incorporate a realistic physical model of the acquisition process can

improve spatial resolution [311]. Lower resolution, higher sensitivity collimators are preferable for brain and cardiac studies [312,313].

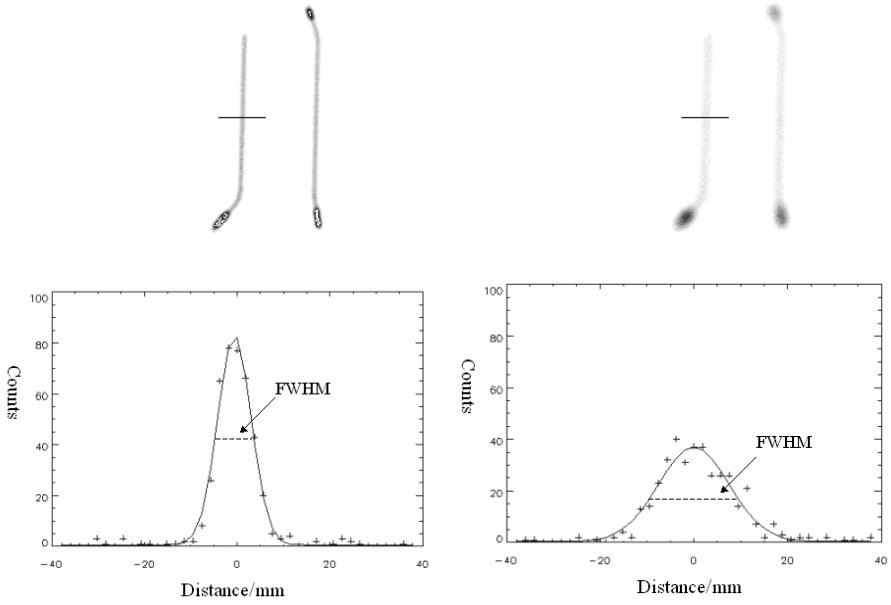


Figure 5.7: Images of line sources at 90 mm (top left) and 290mm(top right) source detector distance. Siemens Orbitor camera was used with a general purpose collimator. Image blurring worsens the farther the line source is from the camera. The horizontal solid lines indicate where the line profiles were measured. The line profiles for 90 mm (bottom left) and 290 mm (bottom right) are shown. The system resolution at 90 mm and 290 mm was measured as FWHM (full width at half maximum), indicated in the profiles by the broken line. A Gaussian curve (solid line) is fitted to the measured data (+) to estimate the PSF, and to compute the FWHM.

5.4.2 Attenuation

Because of interactions with body tissue, a significant number of photons are absorbed or deflected from their original path so that they will not be detected by the camera (ray 5, Figure 5.2). A detailed discussion of the impact of attenuation and its compensation on cardiac SPET was given by King *et al.* [35,36]. The removal of these photons from the beam is referred to as attenuation. For a monochromatic photon beam, attenuation is given by the equation below,

$$I = I_o e^{-\mu_i \delta x_i}, \qquad (5.4)$$

where I_o is the incident intensity, I is the intensity transmitted through a small volume of tissue of thickness δx_i and uniform attenuation coefficient μ_i. Table 5.2 lists the linear attenuation coefficient for several types of body tissues [35]. The fraction of photons attenuated from a narrow beam after passing 5 cm water-equivalent tissue is about 46% for 140 keV photons (μ = 0.15 cm^{-1}) emitted by ^{99m}Tc. In the case of SPET imaging, broad beam conditions apply, accounting for photon loss due to both absorption and

scatter [285]. Furthermore, photons from regions not directly exposed to the beam may be scattered with an increase in detected count-rate above what would otherwise be expected. For 140 keV photons in water-equivalent tissue, the μ value is approximately $0.12cm^{-1}$ for broad beam geometries (thoracic or abdominal imaging) compared to $0.15cm^{-1}$ in narrow beam condition [284]. In a typical study, photons emitted from the center of the abdomen will be reduced by a factor of 5 compared with emission in air. Neglecting this effect will lead to an underestimate of the true radionuclides concentration, which is most pronounced at the center of the body.

Material	Density / kgm^{-3}	Attenuation coefficient at photon energy 73KeV /cm^{-1}	140KeV /cm^{-1}
Air	1.3	0.00	0.00
Muscle	1000.0	0.19	0.15
Lung	330.0	0.06	0.05
Bone	1850.0	0.43	0.29

Table 5.2: The table lists the density and linear attenuation coefficients for several body tissues at 73 keV and 140 keV corresponding to ^{201}Tl and ^{99m}Tc respectively. Air is included for comparison.

The fraction of photon transmitted through the body will vary with the tissue traversed as a result of the differences in attenuation coefficient for different tissues (Table 5.2). Algorithms have been suggested for attenuation correction in SPET [284,314,315]. However, with analytical attenuation correction, reconstructed images result in a varying noise level across the image because the correction factor, such as the Chang algorithm [314], depends on spatial distance from the image center. Attenuation correction usually relies on the availability of an accurate, patient and energy-specific distribution of attenuation coefficient values (attenuation map). The attenuation map is estimated with a transmission (TR) scan (Figure 5.8) by using a line source or in newer systems, a combined CT/SPET scan, and then taking the ratio of the transmitted intensity to the incident intensity [316,317]. Equation (5.4) can be rearranged to solve for μ_i as follows:

$$\ln\left(\frac{I_o}{I}\right) = \sum_{i=1}^{n} \mu_i \delta x_i \ . \qquad (5.5)$$

An estimate of the attenuation map is obtained by reconstruction (Section 5.3) from the attenuation profiles given by Equation (5.5). Figure 1.4 (Chapter 1) shows an example of the intensity profile and its corresponding attenuation profile. The TR scans are similar to CT (Subsection 1.2.1, Chapter 1) but use a higher energy radiation and

different detector configuration, resulting in limited spatial resolution and less soft-tissue detail than CT.

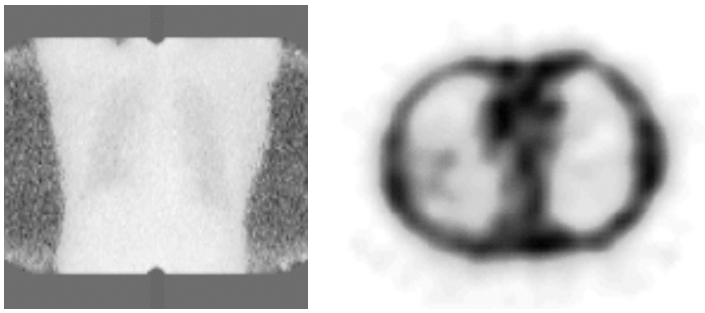


Figure 5.8: One of a number of transmission scans (left) used in reconstructing the attenuation map (right). Note the outlines of the lungs in the transmission scan.

5.4.3 Scatter

When a photon interacts with matter, its energy and direction may change (ray 3, Figure 5.2). Besides primary detected photons, the camera may also detect photons that have interacted with the body. If photons are coherently scattered, no energy is lost and, so long as the scattered direction falls within the admittance cone of the collimator, there is no way to distinguish these photons from primary photons. In nuclear medicine, incoherent scatter (energy lost) is far more likely than coherent scatter. When photons lose energy and change direction due to incoherent (Compton) scatter, their energy could still fall within the photopeak energy window due to the finite energy resolution of the system. In Compton scatter, a gamma photon interacts with an orbital electron, losing energy and changing its direction according to the following equation,

$$E'_{\gamma} = \frac{E_{\gamma}}{1 + \frac{E_{\gamma}}{511}(1 - \cos\theta)}.$$

where E_{γ} is the energy of the original photon, E'_{γ} is the energy of the scattered photon, both in units of keV. The scatter angle θ measures the deviation of the scattered photon from its original direction and takes a value of 0° to 180°. The loss of energy corresponding to 30° scatter is only about 5 keV for ^{99m}Tc 140 keV photons [303]. Since the energy resolution of the detector is about 10%, many of the detected photons carry incorrect position information. A typical energy spectrum recorded from an Anger camera is shown in Figure 5.9. The primary photons are recorded as a broad photopeak. The breadth of the photopeak will include many scattered events, which degrade the SPET image. Scatter accounts for approximately 40% and 50% of the total counts

acquired in a SPET study with ^{99m}Tc and ^{201}Tl, respectively [318]. Effects of photon scatter include a decrease in contrast, loss of resolution and an increased background count. A detailed discussion of scatter and its correction can be found in [284,285,319,320,321,322,323].

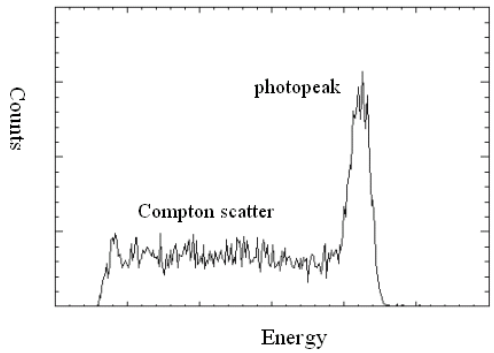


Figure 5.9: Schematic diagram to illustrate an energy spectrum of a typical measurement on a gamma camera measured by the pulse-height analyzer. The photopeak corresponds mainly to the primary photons while the plateau is mainly due to the Compton scatter.

5.4.4 Patient motion

It is well recognized that patient motion during image acquisition can cause artefacts that can compromise image quality, accurate quantification and clinical interpretation. Because of a long acquisition time in SPET (typically > 20 minutes), patient motion is virtually unavoidable. Motion may be extrinsic, such as head movement, or intrinsic, such as the beating of the heart. Movement of patient causes either blurring of projections or inconsistencies in the projection data and tends to produce artefacts in the reconstructed SPET image. In cardiac imaging, the change in depth of respiration may change the position of the heart in the course of data acquisition. Germano [303] reported that 10% to 20% of scans are affected by artefacts due to patient or organ motion in the cardiac SPET, and motion greater than 2 or more pixels results in clinically significant artifacts [90]. Myocardial motion often involves upward creep, bounce and return or nonreturning shift of the heart [303]. The deformation of the myocardium and respiratory-induced motion complicate the task of correction for myocardial motion artefacts. Brain SPET is also susceptible to movement artefacts.

Several SPET motion correction algorithms have been proposed. One approach is to track the motion of the object of interest and apply the measured motion to transform the projection images [324,325]. Since tracking the motion by direct measurement is difficult to implement, attempts to use the measured projection data have also been

proposed [91,92]. The proposed motion correction method is based on a fully 3D algorithm to estimate the patient orientation which best matched each reconstructed projection with its corresponding measured projection.

5.4.5 Partial volume effect

The Anger camera has a characteristic "sensitive volume" determined by its resolution. When an object partially occupies the sensitive volume, there is an apparent decrease of measured signal or counts. The effect is known as the partial volume effect. Rosenthal *et al.* [284] suggested that partial volume effect is due to the mismatch of system resolution, object size and pixel size. One consequence of this effect is that objects of identical concentration are seen to decrease in intensity with decreasing size. As a result, lesion contrast may be reduced and difficult to distinguish from background. Since the total counts are preserved for the smaller object, they are distributed over an area larger than the physical size of the object. Thus the smaller object appears to be larger and has lower average counts than would be expected. The partial volume effect is important for both qualitative and quantitative interpretation of SPET images. The effect makes it difficult to measure the absolute activity using SPET, if objects of different size are imaged, the relative values may be misleading. For example, small central brain structures may appear to have fewer counts than the brain cortex. Therefore, brain perfusion will appear reduced in these central brain structures compared to the cortex. As the dimension of the object becomes larger with respect to the resolution of the detector, the detected maximum counts will approach the true counts emitted by the object. Several studies describe methods to correct partial volume effect using registered anatomical images [326,327]. Koole *et al.* [328] compared several strategies to reduce the partial volume effect with the aim of accurately measuring the true tissue tracer activity within a tissue compartment.

5.4.6 Noise

Due to the limited number of detected photons, there is a substantial amount of Poisson noise in the projection data. The noise is assumed to be spatially uncorrelated, i.e. the statistical variation of signal in any one pixel is independent of the fluctuation in any other pixel. Based on the Poisson statistics, the standard variation of counts in any pixel is estimated by the square root of the average value of counts in the same pixel. The

signal-to-noise ratio (SNR) is approximated by $N/\sqrt{N} = \sqrt{N}$. Thus sufficient counts are often required per projection for acceptable reconstructed image quality [329,330,331]. However, the nature of noise in the SPET image, reconstructed from the Poisson noise dominated projection, is no longer governed by Poisson distribution. A more detailed discussion of the noise properties of reconstructed image, and its dependence on the image reconstruction algorithm, will be given in the next Section.

5.4.7 Noise properties of FBP and OSEM

Several studies of the noise properties of FBP and MLEM reconstruction have been reported [302,332,333,334,335,336,337,338]. In FBP, the ramp filter amplifies high frequency noise. This can be limited by a low-pass filter [301]. The lower the cut-off frequency of the filter, the greater the degree of smoothing (Figure 5.10 top row). Smoothing reduces noise amplitude but also image resolution. At larger cut-off frequencies, a sharper image is observed but its streak artefact is more distinct. For FBP, it was demonstrated that noise is correlated over relatively long distances and its correlation function is characterized by negative side-lobes as the projections may contain negative values after filtering [339]. This effect leads to the image noise spreading from the high counts region to the low counts region and results in a relative constant noise magnitude of globular structure throughout the image. MLEM reconstruction does not suffer from these limitations and demonstrates improved noise properties. Noise amplitude is signal-dependent in MLEM images, producing less noise in low count regions [334,335,338]. The streak artefacts, common in FBP reconstructions, do not appear in the MLEM reconstruction (Figure 5.10). They are more pronounced in FBP reconstructions that contain regions of locally high activity. For example in bone SPET studies, the streak artefact observed has been shown to be intense with FBP but strongly reduced when OSEM is used [340]. Activity is not spread outside the body so that the body outline is more clearly defined in OSEM images than in FBP images. These differences in appearance may have important implications for the performance of image registration, particularly in terms of the ability to align intermodality images. OSEM speeds up the MLEM reconstruction and accelerates its convergence. However, unregularized MLEM reconstruction exhibits noise deterioration with increasing iterations. Note the increased image noise after 4 iterations in Figure 5.10. From the point of view of image registration, one of the limiting factors

is the image quality, and the most important element of image quality is the noise level. In OSEM, once the number of subsets is chosen, the number of iterations entirely controls the final noise of the reconstructed images. For FBP, the principal mechanism for controlling image noise is the cutoff frequency setting of the low-pass filter.

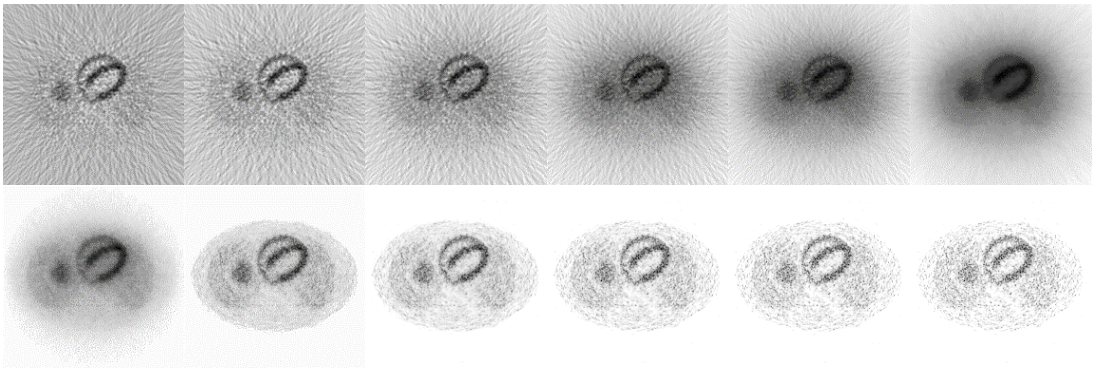


Figure 5.10: Reconstructions from 128 projections over 180^{o}, (with Poisson noise but without scatter and non-uniform attenuation) using FBP (top) and OSEM (bottom). All images are displayed with reversed contrast and are individually scaled. FBP images show the effect of decreasing (left to right) the cut-off frequency of the low-pass filter. The lower the cut-off frequency, the smoother the reconstruction. At higher cut-off frequencies, streak artefacts become more distinct The images in the bottom row show OSEM reconstruction using 4 ordered subsets and 1, 4, 8, 12, 16 and 20 iterations, from left to right.

5.5 The SPET torso phantom

The ability of MLEM reconstruction to incorporate the imaging physics into the reconstruction algorithm makes it very attractive. Improvements in computer technology and the development of the ordered subsets algorithm make the MLEM reconstruction a clinically practical alternative to the FBP reconstruction. Iterative OSEM reconstruction has been shown to give superior noise characteristics to FBP reconstruction [340]. Noise properties of FBP and OSEM reconstructions have been shown to affect image quality [341,342]. However, the difference between OSEM and FBP in the performance of multimodality image registration remains an open question. One of the aims of this study was to determine whether the OSEM reconstruction performs better than classical FBP reconstruction in SPET-CT registration. The performance of multimodality image registration is evaluated using a realistic digitized torso phantom containing a thoracic cavity, liver, heart, anatomically accurate bones and other soft tissues. Furthermore, images of the phantom are reconstructed by FBP and OSEM to investigate the effects of the registration algorithm on registration performance for SPET-CT registration under different conditions

5.5.1 3D digital phantom

Most digital phantoms are usually designed for a specific purpose such as to simulate uptake in the brain (e.g. digital brain phantom in [259,252]) and mimic cardiac SPET imaging (e.g. MCAT phantom [37]). Ordinary MR or CT image of a real subject correlated with post mortem sections can also be used to synthesize a SPET dataset (e.g. Zubal phantom [249,250] and the Visible Human data set [251]). Since anatomically segmentation is not available for the Visible Human data set and the MCAT phantom is not an ordinary medical image, the SPET simulation described here is based on the Zubal digital phantom. The phantom set consists of a CT image and an anatomically segmentated image. The CT volume array represents a high-resolution model of the human anatomy. Segmented Zubal images (to be referred to as the phantom) were created from the CT images by manual segmentation (data supplied with the Zubal phantom). The original CT images were acquired in a 512×512 matrix with a resolution of 1 mm in the transverse plane. The axial resolution is 10 mm from the neck to the mid-thigh and 5 mm from the neck to the crown of the head. Since the computational cost of registration increases with the size of image matrix, the CT and the phantom were resliced from 512×512×97 matrix (with 1×1×10 mm^3 voxels, 25 Mbytes memory) to 128×128×243 matrix (with 4×4×4 mm^3 voxels, 4 Mbytes memory) by downsampling using trilinear interpolation. The segmentation distinguishes 57 body organs (Appendix). It gives a realistic model of anatomical structure that provides spatial bounds on the distribution of the radioactivity and the attenuation coefficient in each organ.

The activity and attenuation coefficient values were assigned to individual organs in the digital phantom based on either published data [343] or clinical cases. To estimate the activity distribution from clinical studies, SPET images were rigidly registered to the Zubal CT image. The registration accuracy was unknown at this point but provided a means of identifying specific organs. The source distributions were estimated by applying the anatomical segmentation to the registered SPECT and estimating the mean counts in each organ. Based on the mean counts, activity distributions were assigned to each organ of the digital phantom. Thus organs and soft tissues are allocated different uptakes to mimic the activity distribution of a specific radionuclide. In addition, to provide a realistic non-uniform attenuation coefficient distribution, attenuation coefficients were assigned according to the tissue type and emission energy of a particular radionuclide. Figure 5.11 shows selected transaxial slices of the SPET

phantom and the corresponding attenuation maps. Based on the labels in the anatomical segmentation, an energy-specific 3D attenuation map and radioactivity distribution can be modelled to simulate clinically realistic image data. The attenuation map used in this study represents an ideal case of a perfect attenuation for torso simulation. In practice, patient-specific attenuation maps would be obtained using simultaneous transmission scans. Such maps often contain noise and reconstruction artefacts (Figure 5.4). The phantom was further downsampled, with interpolation, to a 64×64×121 matrix. The final voxel was a cube of size 8 mm. Two SPET studies were used for simulation purposes, namely ^{67}Ga citrate and ^{99m}Tc-labelled MDP. Imaging with ^{67}Ga is particular useful for assessing both Hodgkin's and non-Hodgkin's lymphoma and detection of inflammation [330,331] whereas MDP is commonly used for bone imaging. Bone SPET is important for detection of a wide range of abnormalities in the skeleton and is often helpful in the management of patients with neoplastic disease [329].

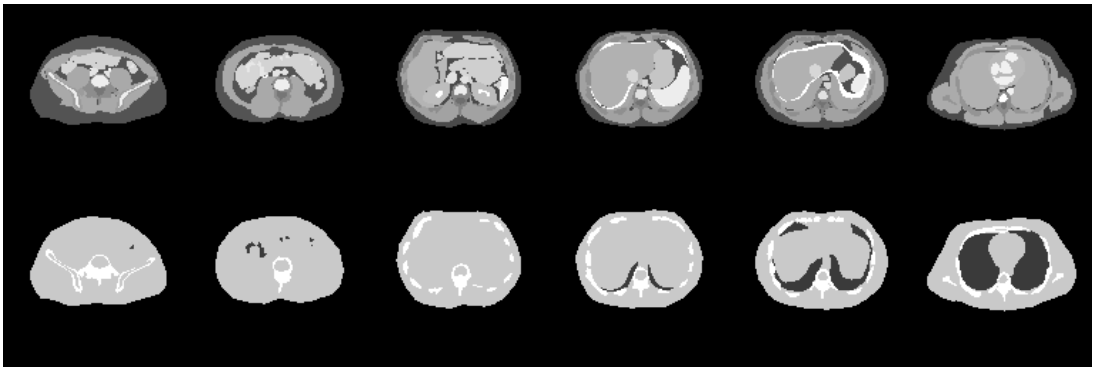


Figure 5.11: Six transaxial slices selected from the 243 slices of the 3D Zubal phantom showing anatomical segmentation (top row) and the corresponding attenuation map (bottom row). Each image is individually scaled.

5.5.2 Projection data

A set of 2D projections was generated from the phantom with a projector, which modeled a medium-energy general-purpose parallel-hole collimator and the effects of attenuation, with system resolution at a distance from the collimator based on manufacturer specifications. Projection data were calculated based on Larkin's implementation on MLEM for ordered subsets reconstruction algorithm [98]. The projection data g_i is related to the image f by $g_i = \sum_{j=1}^{D} a_{ij} f_j$ where weight a_{ij} represents the probability accounting for the i^{th} detector counts in the camera from voxel j in the image. In order to evaluate the transition matrix (a_{ij}), three look up tables are prepared

to account for (i) geometry, (ii) attenuation and (iii) camera blurring. The geometry describes the projection ray position for each pixel at each projection angle. Assuming parallel bin geometry, each voxel's perpendicular distance to the camera (source-camera) was also calculated. Since these distances are the same for all slices of the image, computation for a single slice is sufficient. With an $n \times n$ acquisition matrix for the detector and n projections, there are n^3 pieces of information in the geometry lookup table. The photon intensity emitted from an image voxel to the detector suffers attenuation. The attenuation coefficients are calculated along the ray normal to the detector from each voxel in the image. The attenuation coefficients scale the image values according to (Eq. 5.4). With n^3 voxels in the image and n projections, the size of lookup table for attenuation coefficient is n^4. Assuming a linear loss of resolution with source-detector distance d, the point spread function (PSF) of the detector describes the spread of the counts about the geometric projection point by $\exp(\frac{\sigma^2}{-2t^2})$, where t is the offset from the center of the geometric projection, $\sigma = \alpha + \beta \times d$ is defined by the collimator constant α and the collimator scale β. The spread is used to account for the camera blur based on a normal distribution. The distribution is modeled as a Gaussian and a lookup table for each source-detector distance d and offset t is produced.

In the ^{99m}Tc-MDP SPET simulation, $\alpha = 1.45$, $\beta = 0.0174\,\text{mm}^{-1}$ and d is the source-detector distance in mm. Therefore the system resolution at 100 mm was 7.53 mm ($\text{FWHM} = 2.36\sigma$). For the purpose of projection and reconstruction, assigning tissue-specific attenuation coefficients within the phantom created an attenuation map. Broad beam attenuation coefficients of 0.12, 0.04 and 0.215 cm^{-1} for 140 keV photons were employed for soft tissue, lung and bone, respectively [36,344]. In ^{67}Ga SPET simulation, the collimator parameters were set to $\alpha = 2.15$ and $\beta = 0.0275\,\text{mm}^{-1}$. This gave a system resolution of 11.56 mm (FWHM) at 100 mm. Attenuation coefficients were set to a weighted of attenuation coefficients at 93 keV and 185 keV, with the weights reflecting relative emission abundance and the camera efficiency [345]. The 300 keV and higher energy emissions were excluded, as they are not usually acquired clinically. Based on the phantom, an energy specific attenuation map was created. Attenuation coefficients assigned to soft tissue, lung and bone were 0.134, 0.045 and 0.24 cm^{-1}, respectively. Again, broad beam geometry was assumed.

According to Larkin's implementation on MLEM [98], given a pixel and a projection angle, the projection matrix is produced by look up the projection bin, its offset and depth from the geometry table to compute the PSF that give the matrix. Finally, this matrix is scaled by the lookup attenuation coefficient in the table for each projection angle and pixel. Poisson noise is added to the original noise free projection pixel. SPET is simulated from the phantom using a 64×64 pixel grid with 8×8 mm pixels. Over the 360º span, 64 equally spaced projections were generated with the radius of rotation of 250 mm. Non-uniform attenuation and depth dependent resolution are included in the projection to a mimic real clinical data acquisition situation, but scatter was not included in the simulation. This corresponds to ideal scatter rejection or complete correction [342]. Figure 5.12 shows typical projections for ^{99m}Tc-MDP (MDP, for short) and ^{67}Ga SPET.

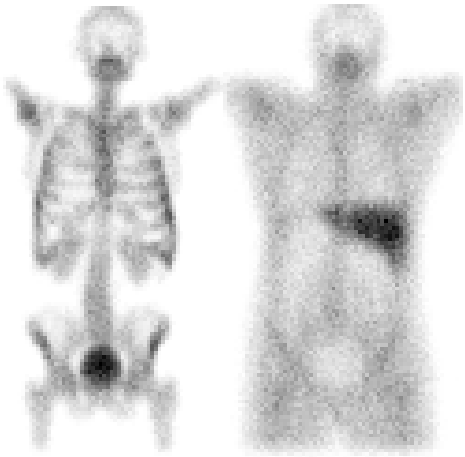


Figure 5.12: A typical planar projection image for MDP (left) and gallium (right) SPET shown with reversed contrast.

5.5.3 Image reconstruction with FBP and OSEM

The reconstruction algorithms considered in this study are FBP and OSEM. Each algorithm is applied to the same set of noisy projection data. Each SPET slice was reconstructed into a 64×64 pixel array that was 1 pixel (8 mm) thick. In the case of OSEM, attenuation was modeled directly in the projector and the back projector, using subset sizes 4 (OS4) and subset size 8 (OS8), with/without non-homogeneous attenuation compensation and no depth-dependent resolution compensation. Given 64 equally spaced projections over the 360º span, one iteration of OS4 corresponds to 16 iterations of MLEM. For FBP, the 3D projections of the image data were transferred to an ADAC Pegasus workstation for analytical reconstruction based on the built-in filtered back-projection (FBP) algorithm. Post-reconstruction filtering was performed

using a 3D Butterworth filter of order 5 at various cutoff frequencies, with and without Chang's attenuation correction.

5.6 Conclusion

In this chapter, two popular approaches, namely FBP and MLEM, for SPET reconstruction are introduced. FBP is a fast analytical reconstruction algorithm based on a simple assumption that the acquired data are simple projection ray sums. MLEM is an iterative method attempted to model the acquisition process. It is then followed by the discussion on image degradation factors in SPET. These include collimator factors (e.g. depth dependent resolution), patient factors (e.g. motion) and factors arising from interactions of radiation with tissues (e.g. scatter and attenuation). These factors, together with the reconstruction algorithms, influence the quality of reconstructed SPET images. For example, when attenuation, depth dependent resolution, scatter and Poisson noise are present, FBP approaches are likely to produce artifacts (e.g. streak artifacts) and inferior noise characteristics, which is a major drawback in SPET reconstruction. The iterative MLEM algorithm has a number of advantages over FBP. It improves the estimate of the image by incorporate a realistic model of the acquisition process that can account for noise, collimator geometry, non-uniform attenuation and scatter. MLEM is computationally expensive but a fast version of the algorithm based on ordered subsets (OSEM) and the improvements in computing technology (faster processor speeds), have made OSEM a popular choice in SPET reconstruction. The superiority of noise properties in OSEM compared to FBP is well documented but its effect on image registration, especially for SPET and CT, is unclear. The use of a digital phantom to simulate projection data degraded by attenuation and depth dependent resolution is included in the projection data, but only attenuation compensation is applied in the reconstruction. The effects of the reconstruction algorithms on registration performance for SPET-CT registration under different settings are uncertain, which will be investigated further in Chapter 6.

Chapter 6

Analysis of registration in non-brain SPET

Relatively few reports describe multimodality image registration outside the brain, especially in the thoraco-abdominal region (Section 5.1) and the performance of registration in this region has not been adequately analysed. It is of particular interest to determine how the new, voxel-based, similarity measures perform in this region. Over the past decade, registration algorithms have evolved from feature-based approaches to voxel-based methods [47,49, 52,56,99,346]. Unlike the feature-based methods, which require features to be extracted from the images prior to registration, the voxel-based approaches generally use the full image volume and operate directly on intensity values. These registrations have been shown to be accurate [122,123,183,242] and have been applied extensively in the brain for multimodality image registration with success. The experiments described in this chapter aim to extend our knowledge of the performance of voxel-based registration methods to the non-brain region.

To facilitate the study of SPET-CT image registration in the thoracic-abdominal region, a digital phantom has been constructed (Section 5.5). Using the phantom, a patient-like source distribution was simulated, based on clinical technetium-99m labeled MDP and gallium-67 citrate SPET studies. The activity distribution in specific organs was obtained from the clinical SPET image, following a pre-registration to an anatomically segmented CT image. The data were forward projected and reconstructed using OSEM. Attenuation and noise were incorporated in each projection. Random misalignments, each with a 6-parameter rigid-body transformation, were applied independently to the CT image. The SPET image was registered rigidly to each transformed CT image. The registration accuracy was estimated from the residual displacement of a vertex of a hypothetical box drawn around the phantom, averaged over the eight vertices.

6.1 Introduction

The majority of the published work on anatomical and functional image registration is focused on the brain, with few studies of the thorax, abdomen, pelvis, and other non-brain areas. Nevertheless, various registration approaches for multimodality image registration in the non-brain region have been reported in literature. For example, the

sum of absolute differences method was used to register thoracic SPET and PET image [165]. External fiducial markers [75,76,79,80] and internal anatomical markers [186,282] have been suggested to register SPET and CT images of the abdomen. Other groups reported the use of mutual information [186, 75, 248], "head and hat" algorithm [78,145], dual-isotope scans [81,130], transmission map [260] and Compton scatter [347] to registered SPET to CT/MR in the non-brain region. Most of these studies were focused mainly on the performance of a particular similarity measure without addressing the factors arising from the choice of SPET reconstruction method. Filtered back projection (FBP) was the method of choice in the reconstruction of SPET data until recently. However, an accelerated expectation maximization algorithm based on the use of ordered subsets (OSEM) is increasingly used in place of FBP. In general, MLEM images have better noise characteristics than FBP images in the lower count region. Without attenuation correction, noise is constant across FBP reconstructed images. Low count regions in FBP images may produce high noise. In contrast to FBP, noise in MLEM is related to signal, and MLEM reconstruction generates less noise in low count regions [38]. Compared to FBP, streak artefacts do not appear in MLEM reconstructions (Subsection 5.4.7). Hence, MLEM leads to smaller intensity errors in the joint intensity histogram and may argue to be better at achieving higher registration accuracy. The effect of the choice of the reconstruction algorithm on the registration performance has not been reported. In particular, it is interesting to note that only limited studies, if any, have been reported to determine if registration accuracy is dependent on image noise level.

Performance of multimodality image registration has been studied extensively in the brain region for various similarity measures [122,123]. It is unclear how similarity measures, such as MI, NMI and SCR, perform in non-brain regions. Do these measures produce similar registration accuracy as in the brain region? Does the performance of SPET-CT registration depend on the choice of the reconstruction method (FBP or OSEM)? At the heart of these similarity measures (MI, NMI and SCR) lies a joint intensity histogram, which is needed to estimate the joint probability and marginal probability of images. In these methods, the voxel intensities are typically requantized first. The number of bins and other factors that affect the joint intensity histogram may also influence the registration accuracy. Thus, in a comprehensive analysis of registration performance, it is imperative to determine the extent of the influence of pre-

registration processing, including intensity windowing, attenuation compensation, and resolution matching. Registration of the thorax and abdomen presents a somewhat different problem from that of the brain, due to the elasticity and movement of organs. Due to different protocols and fields of view in different modalities, the imaged volumes commonly overlap only in part. Another difficulty is the poor definition of organ and body boundaries in SPET. In clinical CT, the field of view is usually kept small to minimize the radiation dose to the patient. In SPET, the field of view is limited by the geometry of the gamma camera. CT and SPET images to be registered may therefore represent different, partially overlapping, volumes. The effect of the presence or absence of the body boundary in SPET and the issue of noncoincident image volumes has received scant attention. Therefore, one purpose of this study is to examine the effect of body boundary and field of view on registration.

The limited size of the existing physical phantom makes it difficult to vary the field of view and study its effect on registration. Issues of versatility and ability to control individual factors to model acquisition of data; therefore, physical phantoms were not used in this study. In order to address the above problems, a digital phantom (Section 5.5) was used to investigate SPET-CT registration in the thoracic-abdominal region. Specifically, the purpose of this study was to investigate: (1) the relative performance of the three similarity measures (SCR, NMI and MI), (2) the influence of the choice of reconstruction algorithm (OSEM or FBP), with and without attenuation correction; and (3) the effect of image attributes, such as intensity window settings in CT, relative image resolution, relative image volume and degree of overlap between SPET and CT images as well as the presence of boundary information in SPET. Registration performance was evaluated by examining how well the corresponding SPET data had been registered to a transformed CT under various conditions. The radionuclide distribution simulated clinical MDP and gallium-67 SPET studies.

The rest of this chapter is arranged as follows. The implementation of phantom studies is described in Section 6.2. The evaluation methods used for examining the registration are presented in Subsection 6.2.6 and the experiments are in Section 6.3. Experimental results are presented in Section 6.4 and discussed in Section 6.5.

6.2 Methodology

The simulations of radionuclide distribution in clinical SPET studies, specifically methylenediphosponate (MDP) and gallium-67 (^{67}Ga), have been described in Section 5.5. Brief descriptions are given below. For the purpose of registration, the simulated SPET image served as the floating image while the CT image was the target image.

6.2.1 Image data

To study multimodality registration in the thorax and abdomen, the Zubal digital phantom [250] was used to generate an image data set as described in Subsection 5.5.1. Slices 74-185 were extracted from the CT volume (128×128×243 matrix with 4×4×4 mm^3 voxels), representing the thoracic-abdominal region. The extracted volume was 57% of the original Zubal CT volume. In SPET, the whole image, from the top of the head to the mid-thigh, was used as the default image.

6.2.2 Projection data

Projections were simulated as described in Subsection 5.5.2. The planar projection counts and noise levels were chosen based on clinical studies. Total counts for a typical planar projection were approximately 198k for ^{99m}Tc-MDP and 208k for ^{67}Ga corresponding to the numbers obtained clinically, and referred to as ×1 clinical count level. For ×0.5 count level, the planar projection was approximately half to that of ×1 count level. To distinguish these two cases in the rest of this chapter, image at ×1 is denoted by high-count image while image at ×0.5 is referred by low-count image.

6.2.3 Reconstruction

As described in Subsection 5.5.2, SPET images were reconstructed from the projection data by filtered back-projection (FBP) and ordered subset maximum likelihood-expectation maximization (OSEM). A detailed description of the image reconstruction for ^{99m}Tc-MDP (MDP, for short) and ^{67}Ga SPET is given in Subsection 5.5.3. Briefly, in OSEM reconstruction, 1, 2, 4, 8, 16 and 32 iterations of OS4 and 1 iteration of OS8 were used for SPET reconstruction. For FBP, a 3D Butterworth (order 5) pre-reconstruction filter of cut-off frequency in the range of 0.2-0.7 cycle/cm in steps of 0.1 cycles/cm was investigated. A ramp filter (“infinite” cutoff) was also used.

6.2.4 Similarity measure

Measures based on the co-occurrence of voxel intensity values in the two-dimensional joint intensity histogram: mutual information (MI), normalized mutual information (NMI) and symmetric correlation ratio (SCR), were compared. Requiring only the assumption that there is a probability dependence of PDFs for MI and a conditional variance for SCR between groups of voxels that have similar values, these methods are more robust than feature-based methods and are generally applicable to multimodality image registration. Studies [181,182] have demonstrated that the correlation ratio method gives more accurate results than mutual information when severe sub-sampling is used for multi-levels registration to improve speed comparing to full sampling. A symmetric version of the correlation ratio (SCR) was employed for SPET-CT registration. In the rest of this chapter, "the similarity measures" will refer collectively to SCR, NMI and MI.

6.2.5 Default implementation

The default setting comprised CT intensity windowing, 3-level strategy, and simplex optimization. All studies were carried out on a stand-alone 333 MHz Pentium PC with 192 Mbytes of random access memory.

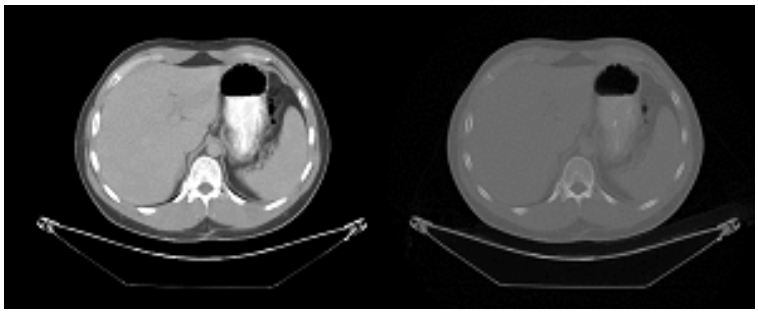


Figure 6.1: CT image with (left) and without (right) intensity windowing.

CT intensity windowing

The intensity histogram of the Zubal CT was found to be largely confined to a narrow range of intensities [800,1250]. In general, specific tissues are visualized on CT with a choice of predefined standard windows optimized for different clinical applications. In this case, however, the aim was to enhance overall contrast among major tissue types while compressing the intensity range to 8 bits. CT intensities were re-mapped linearly from interval [800,1250] to [0,255]. Intensities lower than 800 were set to zero and those higher than 1250 were set to 255 (Figure 6.1). Where intensity windowing was

not applied, the full CT intensity range was rescaled linearly to the available grey scale [0,255] without clipping. In the default setting, intensity windowing was applied in the CT while the SPET images were linearly rescaled to 256 grey levels regardless of the maximum intensity.

3-level strategy

Experimental techniques similar to those described in Section 4.3 were employed. Briefly, a sub-sampling strategy was used to speed up the registration process. A three-level scheme [258] was employed, comprising steps: [8,8,1|24], [8,4,1|32] and [4,4,1|64], where the notation [n_1, n_2, $n_3|b$] was defined in Subsection 4.3.4. For example, a [4,4,1|64] sampling scheme applied to an image of 128^3 voxels will generate 32×32×128 samples for registration. The intensity in each image was rebinned into 64 intensity values. Rebinning was performed implicitly by linear rescaling of image intensity. A tri-linear interpolation was used to obtain the intensity values after image transformation. The downhill simplex optimization method was adopted from Numerical Recipes in C [127]. Optimization similar to that described in Subsection 4.3.5 was used. Note that, for the purpose of the display, all images in this chapter were individually scaled according to their maximum value.

6.2.6 Analysis of registration accuracy

The accuracy of a similarity measure for rigid registration of clinical image data can be established relative to a gold standard, which is based on bone-implanted fiducial markers [122]. The main drawback of this approach is that fiducial markers are often not available for experiment. In the case of synthetically deformed images, one may use the known transformation as the ground truth to evaluate the registration result. The results can be used as a consistency check of performance of the similarity measures under different experimental conditions. The 3D CT image was transformed randomly using rotations (up to ±10° about each of the three orthogonal axes) and translations (up to ±80 mm along each axis). The six parameter values of the rigid-body transformation were drawn at random from a uniform distribution. Registration was then applied to determine rigid registration parameters that best aligned the 64×64 SPET image with the transformed CT image. The registration accuracy was obtained by measuring the

voxel displacement, averaged over eight vertices of a cuboid box drawn around the body surface (Figure 6.2).

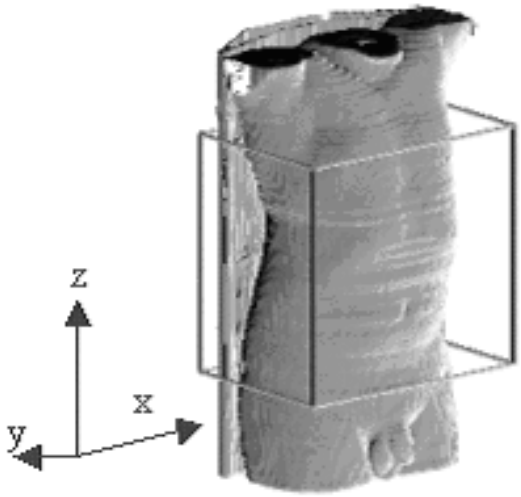


Figure 6.2: The box drawn around the CT volume is used to compute the registration error. Dimensions of the box were (x,y,z)=(46, 31.2, 43.6) cm.

The box will be referred to as the bounding box even though it does not provide a bound along the long axis of the body. The average displacement provided a measure of absolute performance of SPET-CT registration. The misalignment applied to the CT image is denoted by $\theta_{\text{app}}(x)$. A retrospective SPET-CT registration was then applied, producing a recovered transformation $\theta_{\text{rec}}(x)$. In all experiments, the same box was used to compute the registration accuracy even if the image volume was altered. The registration accuracy is defined as in Subsection 4.2.1 [179,242,258],

$$\varepsilon = \frac{1}{8}\sum_{i=1}^{8}\left\|\theta_{\text{rec}}(x_i) - \theta_{\text{app}}(x_i)\right\|. \qquad (6.1)$$

The maximum vertex displacement gives the upper bound on the displacement of all voxels bounded by the box with respect to rotations about an axis through the centre of the box. A rigid-body transformation can be expressed as $\theta(x) = \mathbf{R}x + \mathbf{t}$, where **R** stands for pure rotation and **t** for pure translation. The registration error can be expanded further to look at **R** and **t** more closely,

$$\varepsilon = \frac{1}{8}\sum_{i=1}^{8}\left\|(\mathbf{R}_{\text{app}}\, x_i + \mathbf{t}_{\text{app}}) - (\mathbf{R}_{\text{rec}}\mathrm{x}_{\text{i}} + \mathbf{t}_{\text{rec}})\right\|$$
$$= \frac{1}{8}\sum_{i=1}^{8}\left\|(\mathbf{R}_{\text{app}} - \mathbf{R}_{\text{rec}})\, x_i + (\mathbf{t}_{\text{app}} - \mathbf{t}_{\text{rec}})\right\|$$
$$\le \frac{1}{8}\sum_{i=1}^{8}\left\|(\mathbf{R}_{\text{app}} - \mathbf{R}_{\text{rec}})\, x_i\right\| + \frac{1}{8}\sum_{\text{i=1}}^{8}\left\|(\mathbf{t}_{\text{app}} - \mathbf{t}_{\text{rec}})\right\|, \qquad (6.2)$$

where the suffix distinguishes between the applied and the recovered transformation. It is interesting to examine the mean displacement $\underline{\varepsilon}$ representing the vector sum of residual vertex displacements. Under a rigid body transformation θ, all voxels experience identical translational and angular displacements, and the vector sum reduces to

$$\varepsilon = (\mathbf{R}_{\text{app}} - \mathbf{R}_{\text{rec}})\,\bar{x} + (\mathbf{t}_{\text{app}} - \mathbf{t}_{\text{rec}}), \qquad (6.3)$$

where $\bar{x} = \frac{1}{8}\sum_{i=1}^{8} x_i$ is the mean position of the box vertices, i.e. the center of the box. If the coordinate system is chosen to be anchored to the center of the box, the first term will vanish and the mean displacement vector will be given solely by the residual translation. However, if the error is defined as a scalar quantity (Eq. 6.2), the rotational displacements about the center of the box do not cancel out. Thus the rotational error is given by $\varepsilon_{\text{R}} = \frac{1}{8}\sum_{i=1}^{8}\left\|(\mathbf{R}_{\text{app}} - \mathbf{R}_{\text{rec}})\, x_{\text{i}}\right\|$. The upper bound of the registration error is therefore composed of the rotational error and the translational error $\varepsilon \le \varepsilon_{\text{R}} + \varepsilon_{\text{t}}$, where $\varepsilon_{\text{t}} = \left\|\mathbf{t}_{\text{app}} - \mathbf{t}_{\text{rec}}\right\|$. It implies that the mean vertex displacement error depends on the mean distance of the vertices from the centre of the box. The lower bound of the mean rotational error can be estimated by $\varepsilon_{\text{R}} \approx \varepsilon - \varepsilon_{\text{t}}$ when $\varepsilon_{\text{t}} >> \varepsilon_{\text{R}}$.

The error ε was averaged over independently generated SPET-CT registrations to yield a mean registration error $\bar{\varepsilon}$ for each experiment. Therefore, the upper bound of the mean registration error $\bar{\varepsilon}$ is composed of the mean rotational error $\bar{\varepsilon}_{\text{R}}$ and the mean translational error $\bar{\varepsilon}_{\text{t}}$. It is noted that these errors are not comparable to those of other methods in [122,123] since different parts of body were used to compute the error. Unless otherwise specified, registration performance was evaluated for 50 random transformations of the CT image. SPET images were registered to the transformed CT

images and the transformation recovered. The registration error was averaged over the 50 SPET-CT registrations to give a mean registration error in each of the experiments. The sensitivity of similarity measure to the random error introduced by the uncertainty in defining random transformation in CT is measured by the registration error. The spread of the error (standard deviation) indicates how precise the result of SPET-CT registration is. Unless otherwise specified, two-tailed unpaired t-tests were used as the default method to compare the results statistically of the various experiments. There are cases where gross misregistration occurs, due to registration error where the registration terminates at a local optimum, which is unrelated to the correct alignment. In this case, a statistical test is inappropriate in that it compares means, and means can be strongly biased by outliers. The size of the error for an outlier is not very interesting from the point of view of registration performance (whether it is 150 mm or 175 mm, it is failure for registration). As a default in the rest of this chapter, a registration error of greater than 100 mm is considered a failure in registration, and it is rejected in the computation of summary statistics to prevent possible bias of mean and standard deviation of an error distribution. Empirical results showed that, in some cases, errors ≥ 100 mm were found. So a superior test would be to strip outliers before computing the mean, or to compare medians, and also to count the outliers. It could be that in this case, the medians also will not be significantly different, so the number of outliers is the only differentiating characteristic of registration performance. No observer detected a definite misregistration in the "best" registered image. A threshold of 2.5 pixels (i.e. 20 mm) of the largest voxel dimension (i.e. 8 mm) is used as the threshold of detection of misregistration for SPET-CT registration. In an ideal case, a useful threshold could be constructed for each clinical application, based on the clinical requirements. For the purpose of measuring the robustness of registration with respect to different conditions, 20 mm was set as the threshold to compute robustness. It is measured by the rate of registration having an error greater than the threshold.

6.3 Experiments

6.3.1 Effect of the choice of reconstruction algorithm

In this experiment, the effect of reconstruction algorithms on SPET-CT registration was investigated. OSEM reconstructions were carried out with various levels of iteration cycles and FBP reconstructions were repeated with different values of the cutoff

frequency (Subsection 5.5.3). These two parameters control noise in the reconstruction. A local measure of image noise is given by the standard deviation (SD) of counts in a region normalized by the mean counts and averaged over 3 slices. The region (Figure 6.3), where the image noise was computed, was homogeneous and large enough to ensure sensible statistics. The default setting described in Subsection 6.2.5 and SCR was employed. Independently generated SPET images were generated comprising 100 simulations of ^{99m}Tc-MDP and 100 ^{67}Ga at high clinical count level (×1), each set split equally into 50 simulations with attenuation correction and 50 without. The same experiments were repeated for low-count ^{99m}Tc-MDP and ^{67}Ga images (×0.5). Attenuation correction in SPET imaging is of interest because it removes artefacts that may affect registration. Attenuation effects are greatest in the thoracic-abdominal region because of nonuniformity of attenuation in that region. AC is expected to reduce attenuation artefacts, but it can also change the magnitude of image noise. Much work has been done on attenuation correction in SPET, but its usefulness in image registration remains an open question. In this experiment, the effect of attenuation compensation (Subsection 5.5.3) is also investigated. The mean registration error was again computed as in Subsection 6.2.6 and averaged over 50 registrations corresponding to 50 random transformations of the CT image.

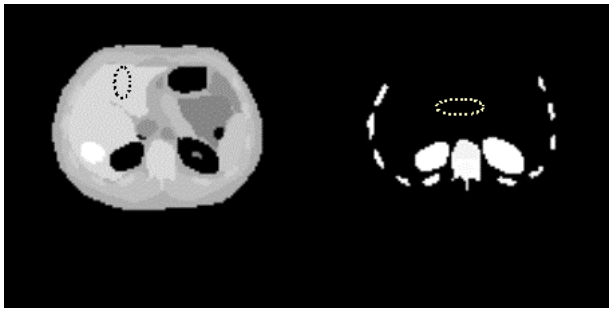


Figure 6.3: Enclosed by dotted contour are the regions in gallium-67 (left) and MDP (right) SPET images where image noise was computed.

6.3.2 Comparison of similarity measures

The purpose of this experiment was to compare the SPET-CT registration performance of SCR, MI and NMI. All 200 SPET reconstructions (Subsection 6.3.1) were used to account for possible effects of the differing reconstruction methods.

6.3.3 Effect of background

In the simulation of SPET scans, scatter was not taken into account. In practice, photon scatter leads to a decrease in contrast, loss of resolution and increased background counts. It is unclear whether scatter rejection will benefit SPET-CT registration. In this investigation, background counts measured in clinical images were used for SPET simulation. Since a background of 7.5% of the liver mean count was observed in clinical studies, this figure relative to the simulated liver mean count was superimposed on the projection data to create background activity (Figure 6.4). The background activity was used to mimic scattering due to septal penetration of the high energy gammas from ^{67}Ga through the collimator as observed in projection data at $\times 1$. This septal penetration will add an approximately uniform background to the image. Only gallium-67 SPET simulation reconstructed with OSEM was used. The registration error was determined separately for each of the three similarity measures: SCR, NMI and MI. This error was averaged over 25 independently generated SPET-CT registrations. The case of zero background activity represented a perfect rejection of scattering in the projection data.

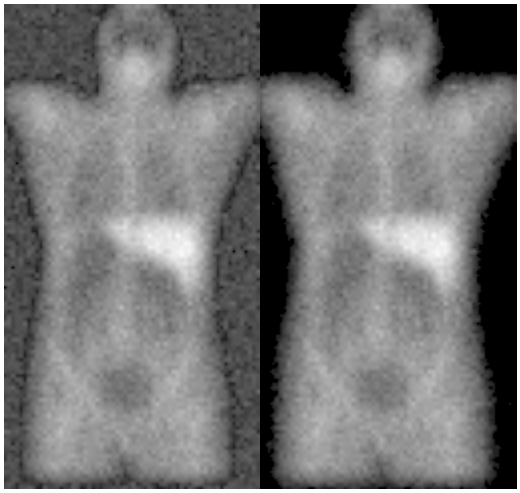


Figure 6.4: A planar projection of gallium-67 SPET (64×121) with (left) and without (right) background activity at high planar projection count levels.

6.3.4 Effect of matching the image resolution

Image resolution in SPET is poorer than in CT and this mismatch may affect the performance of SPET-CT registration. Some studies [258,232] reported filtering of image data to reduce resolution mismatch. The aim of this study was to determine whether matching image resolution between CT and SPET could be used to improve registration. To this end, the CT image was forward projected with the same settings as

for the SPET, with depth dependent resolution and non-uniform attenuation, but without Poisson noise. The reconstruction used one iteration of OS4 with AC. The SPET image was registered, using the default implementation (Subsection 6.2.5), to the blurred CT image, which was transformed randomly. The recovered transformation and the applied transformation were used to compute the registration errors (Subsection 6.2.6). The measurement was repeated over 25 independently generated SPET-CT registrations to yield a mean of the registration error with three similarity measures: SCR, NMI and MI.

6.3.5 Effect of intensity windowing in CT

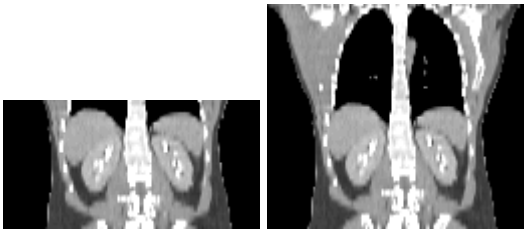
The purpose of this experiment was to investigate whether intensity windowing could improve SPET-CT registration. To assess the effect of CT intensity windowing, registration with and without windowing was performed, with the default setting (Subsection 6.2.5). Twenty five random transformations were applied to the CT and the OSEM reconstructed SPET, ×1 count level, was registered to the transformed CT. In order to account for a possible variation in the effect due to the choice of the similarity measure, all three measures (SCR, NMI and MI) were used. Likewise, the experiment was carried out with both ^{67}Ga and MDP SPET simulations. The registration error was computed as in Subsection 6.2.6.

6.3.6 Effect of body boundary in SPET

The presence of the body boundary in the SPET image provides a spatially extensive feature that is common to both SPET and CT. Given the paucity of common features, it is reasonable to expect that the boundary could improve the registration performance. To assess the effect of missing boundary information on registration, gallium-67 SPET without soft tissue uptake was reconstructed by OSEM with attenuation correction. Because of the streak artifacts outside the body boundary that could affect the registration, FBP was not considered in this experiment. The relative uptake in soft tissue (skin, fat and skeletal muscle) was set to background level (i.e. the mean activity is zero). Gallium SPET was reconstructed at high count (×1) and low count (×0.5) level. The default setting defined in Subsection 6.2.5 was used. Registration error (Subsection 6.2.6) was measured for 25 randomly transformed CT images, with three similarity measures: SCR, NMI and MI, with and without CT intensity windowing.

6.3.7 Effect of absence of lung in CT

In the experiment described in 6.3.6, the effect of the presence of one prominent feature in SPET (body boundary) was investigated. Other prominent features, such as the lung could also provide boundaries to constrain the registration. The purpose of this experiment was to investigate whether registration is affected by the exclusion of the lungs from the field of view. The lungs were largely excluded by extracting only slices 77 to 140 from the original 128×128×243 CT volume of $(4\ mm)^3$ cubic voxels (Figure 6.5). Using the default setting in Subsection 6.2.5, the SPET image was registered to the 25 randomly transformed CT images. The registration was carried out for gallium-67 SPET reconstructed with OSEM and MDP reconstructed with FBP, both at the clinical planar projection count level. All three similarity measures (SCR, NMI and MI) were used.-



CT: slices 77- 140 CT: slices 74 - 185

Figure 6.5: A coronal section through the restricted CT volume that excludes the lungs (left), compared to the volume used in previous studies (right). The range of slice numbers is shown in square brackets.

6.3.8 Effect of field of view

Scanning protocols in CT and SPET produce different fields of views but the effect of such differences on registration is unclear. The purpose of this experiment was to investigate this effect via two situations. In one situation, starting with the same field of view extracted from SPET and CT images, both volumes were progressively reduced concurrently to 79%, 57%, 43%, and 32% of the original value by contraction of the field of view along the long axis of the body. In the other situation, the SPET volume was kept fixed (at 57%) and the CT volume was varied. Based on the default setting in Subsection 6.2.5, the SPET image was registered to 25 randomly transformed CT images and registration errors measured. The investigation was carried out for high count gallium-67 and MDP SPECT using OSEM and FBP reconstructions driven by SCR, NMI and MI similarity measures.

6.3.9 Effect of bounding box size

In the experiments described so far in this chapter, the same bounding box drawn around the Zubal phantom was used to compute the registration accuracy irrespective of the field of view. The Euclidean distance measured for the residual displacement of each vertex, averaged over all eight vertices, yielded the mean registration error. The dependence of the mean registration error on the box size is unclear. Therefore, two experiments were performed. First, the dependence of the mean registration error on the bounding box size was investigated. To this end, 25 random transformations were applied to distort the CT image of default volume of 57% (slices 74-185). Gallium and MDP SPET images reconstructed at the high count level (×1) were registered to the transformed CT image using SCR. Box sizes of 46×31.2×43.6 cm^3 (C_1) and 23×15.6×21.8 cm^3 (C_2, which has half the dimensions of C_1) were used to compute the mean registration error. Second, the effect of using a smaller CT volume (43% instead of 57%) on the mean registration error was studied.

6.4 Results

6.4.1 Effect of the choice of reconstruction algorithms

Variation of the typical mean registration error $\bar{\varepsilon}$ (in mm) in SPET-CT registration with cutoff frequencies for FBP (left column) and iteration number for OSEM (right column), with AC at planar projection count ×1 and ×0.5, is plotted in Figure 6.6. These plots gave broad comparisons across cutoff frequencies for FBP and iterations for OSEM reconstructed images using SCR. With ^{67}Ga SPET, high count images produced better mean registration error than low count images for both of FBP and OSEM irrespectively to cutoff frequencies and iterations. With MDP SPET, mean registration errors for high count and low count images appeared virtually identical, and varied similarly for both FBP and OSEM. The results also show that, with AC, the mean registration error for FBP ^{67}Ga was greater than that for OSEM ^{67}Ga regardless of the cutoff frequencies in FBP and the number of iterations in OSEM (Figure 6.6 first row). Similar observations were obtained for MDP SPET (Figure 6.6 second row).

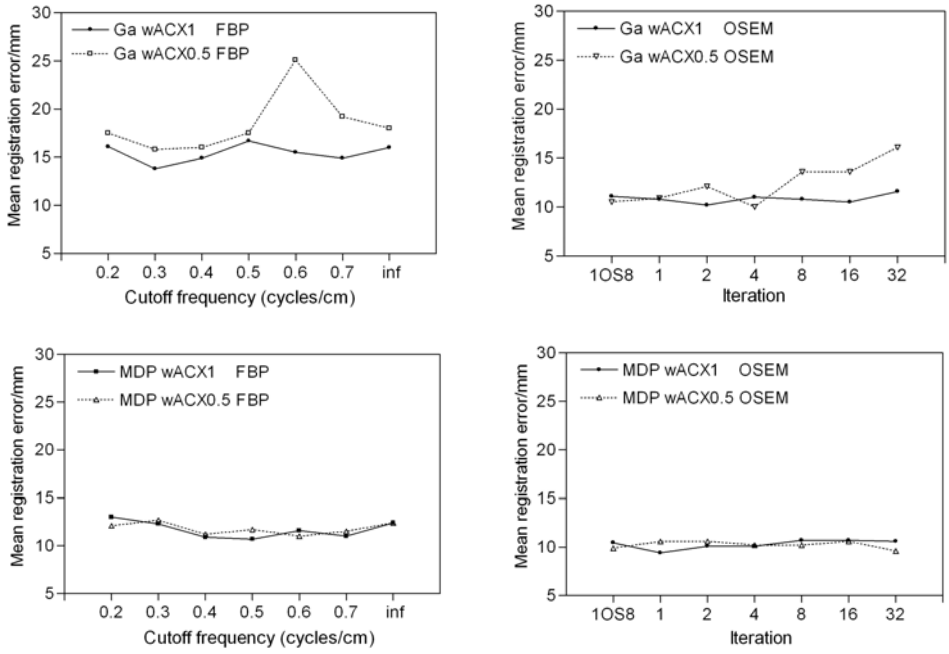


Figure 6.6: Mean registration error (mm) for SPET-CT registration with AC at two image noise levels. Each data point represents the average over 50 registrations. The top row shows registration error for ^{67}Ga SPET. The bottom row shows typical registration for MDP SPET. For FBP (left column), the numbers '0.2-0.7' denote the cutoff frequencies of the Butterworth filter at 0.2-0.7 (in step of 0.1) cycles/cm and 'inf' refers to the cutoff frequency of a ramp filter. For OSEM (right column),'1OS8' refers to 1 iteration of 8 ordered subsets OSEM while the numbers 1,2,4,8,16,32 refer 1,2,4,8,16,32 iterations using OS4, respectively.

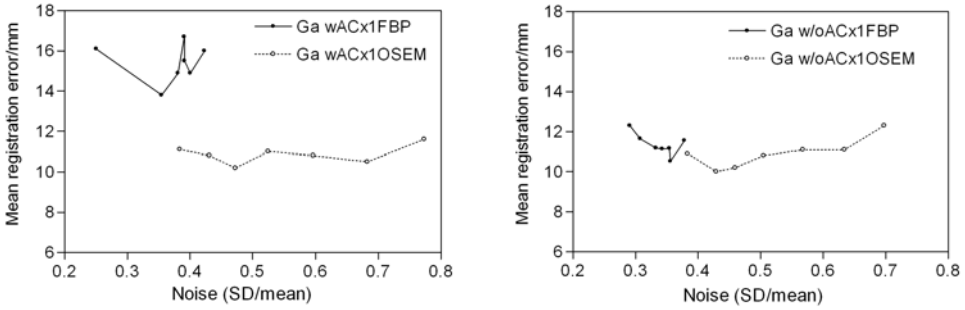


Figure 6.7: Variation of mean registration error (mm) against image noise (SD/mean) for FBP and OSEM reconstructions of gallium-67 SPET for two cases: with (left) and without (right) AC. 'wAC' stands for reconstruction with attenuation correction while 'w/oAC' indicates no AC. Each data point represents an average over 50 registrations.

Since image quality is different for FBP and OSEM, and it varies with reconstruction settings, a comparison of registration performance should be made with reconstructions that are, in some sense, equivalent. One approach might be to select as equivalent reconstructions that are deemed clinically optimal. However, this introduces subjective judgment and does not aid in the understanding of registration performance. The image quality characteristic that is likely to be of greatest significance to registration performance is noise. This suggests the choice of noise-equivalent reconstructions.

Thus finding a noise-equivalent pair requires tuning the FBP filter and the OSEM iteration number until the reconstructed images manifest the same normalized noise measure in a selected region (Figure 6.3). Variation of $\bar{\varepsilon}$ with image noise (SD/mean) is plotted in Figure 6.7 for reconstructions of gallium-67 citrate SPET. The plots show that the range of image noise was different in the FBP and OSEM reconstructions, with a small overlap between the least filtered FBP and least iterated OSEM. The difference in $\bar{\varepsilon}$ between FBP and OSEM is pronounced with AC (about 5 mm). For any image noise where both reconstructions are available, OSEM provides better registration performance than FBP when AC was included in the reconstruction. Similar results were obtained for MDP (not shown here). It is surprising, given the strong averaging, that two FBP (with AC) data points at 0.4 image noise gave quite different results, implying that even though the filter cutoff frequency was changed, the reconstruction quality did not. Statistically, a large difference among sample means implies that the sample size may not be large enough to give a decent estimate of the population mean. To evaluate the effect of FBP and OSEM on SPET-CT registration, images of roughly equivalent noise were used for ^{67}Ga and MDP, with and without AC. Using the two-tailed unpaired t-test, mean registration error for OSEM was found to be significantly smaller (p<0.05) than FBP regardless of image count level of ^{67}Ga and MDP (×1 and ×0.5) when AC was included. For two selected FBP and OSEM implementations that give rise to roughly noise-equivalent reconstructions, distribution of $\bar{\varepsilon}$ is indicated in Figure 6.8. For ^{67}Ga SPET of roughly equivalent image noise, $\bar{\varepsilon}$ (±SD) decreased from 15.5±6.2mm (FBP, f0.6) to 10.8±4.3mm (OSEM, 1OS4) with AC at clinical planar projection count level ×1. The error $\bar{\varepsilon}$ also improved from 16.0±7.3mm (FBP, f0.4) to 10.5±3.0mm (OSEM, 1OS8) with AC at ×0.5. The error bars in the diagram indicate the standard deviation of mean error distribution. Without AC, the difference in the mean registration error between FBP and OSEM was not statistically significant, implying that the difference was significant with AC. To further demonstrate results, ^{67}Ga SPET images with AC for FBP and OSEM at ×1 and ×0.5 are shown in (Figure 6.9). The images look smooth with roughly the same image noise but streak artefacts appears only in FBP images.

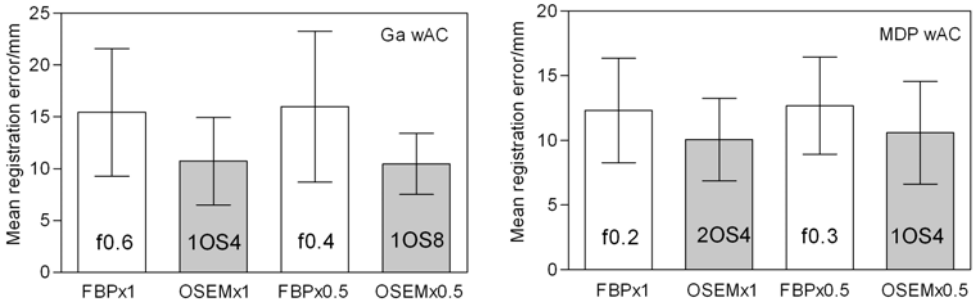


Figure 6.8: Distribution of mean registration error (mm) for ^{67}Ga and MDP SPET-CT registration with AC, where 'f0.6' denotes the cutoff frequency of the Butterworth filter for FBP at 0.6 cycles/cm and 1OS8 refers to 1 iteration of 8 ordered subsets OSEM. Each mean registration error represents an average over 50 registrations. In each bar chart, the first two plots and the last two plots, are noise-equivalent pairs.

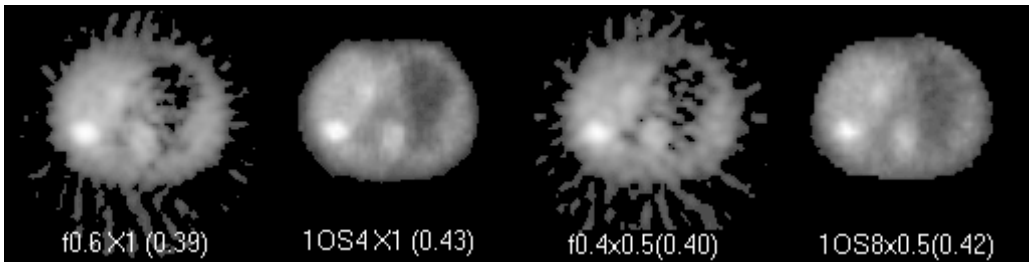


Figure 6.9: A transverse slice showing roughly noise-equivalent reconstructions of gallium-67 SPET. Measured image noise is shown in brackets. the first two pairs and the last two pairs are noise-equivalent. All images are scaled independently.

In general, two options, with and without AC, can be applied to reconstruct SPET image. Again, to make the comparison of registration accuracy meaningful between images at equivalent noise, the mean registration error $\bar{\varepsilon}$ is plotted as a function of the image noise (SD/mean). Separate plots are given for the registration error obtained in the presence or absence of AC, and for each reconstruction technique: FBP and OSEM (Figure 6.10). Two-tailed paired t-test was used to determine the difference in the mean registration error between FBP SPET with and without AC, with pairing asserted by the noise figure. For OSEM, the same test was applied to SPET at roughly equivalent image noise levels. The results reveal that attenuation compensation had little effect on the performance of SPET-CT registration for OSEM-reconstructed SPET. There was no statistical difference ($p < 0.05$) in the mean registration error between OSEM SPET with and without AC. The results in Figure 6.10 (top row) show that the mean error for OSEM SPET (both ^{67}Ga and MDP) was largely unaffected by attenuation correction.

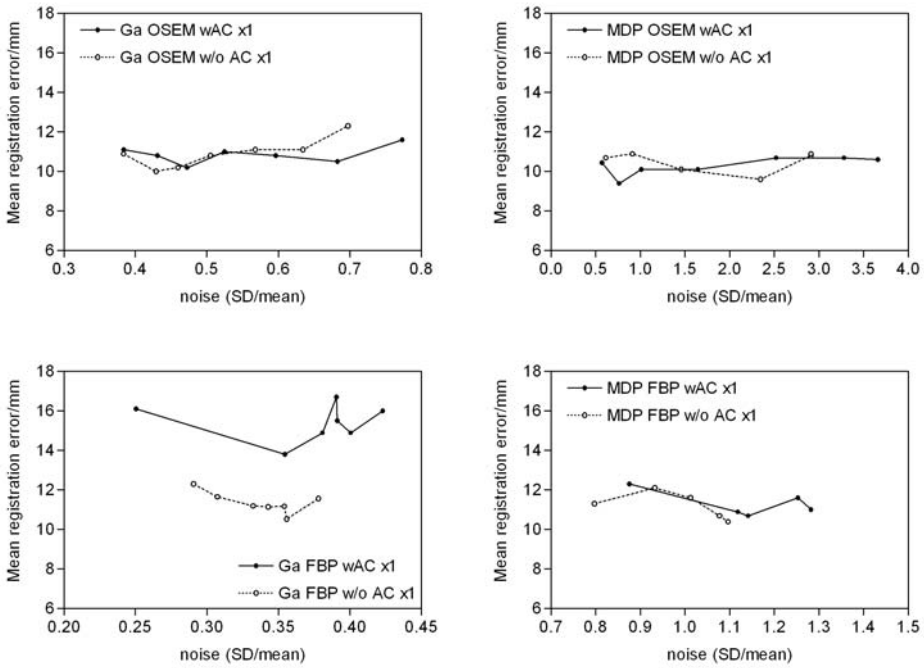


Figure 6.10: The effect of attenuation correction on registration performance for indicated reconstructions. Mean registration error (mm) is plotted as a function of image noise (SD/mean). All registrations used the default implementation with SCR as the similarity measure.

The same outcome was obtained when the image noise level was increased by reducing planar projection count to ×0.5 (not shown). For FBP, AC markedly reduced registration performance for ^{67}Ga SPET but not for MDP (Figure 6.10 bottom row). The difference in the mean registration error was significant ($p < 0.05$) for ^{67}Ga irrespective of image noise level. For MDP, AC did not produce a significant difference in the mean error at either high (×1) or low (×0.5) planar projection count level. Examples of ^{67}Ga SPET at roughly equivalent noise are also shown in Figure 6.11. The first three images look similar except the rightmost image, which is used to show the effect of a ramp filter on reconstruction. The ramp filtered image has a higher noise level but gives better contrast than the first three images without ramp filtering.

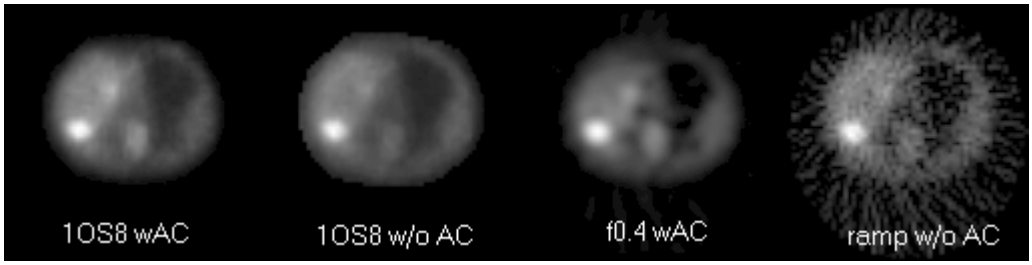


Figure 6.11: Three noise-equivalent reconstructions of a transverse slice of gallium-67 SPET image. From the left: OSEM with 1 iteration of 8 ordered subsets, with and without AC, and FBP reconstruction using Butterworth filter with 0.4 cycles/cm cutoff. For reference, a FBP reconstruction with a ramp filter alone is also shown (rightmost).

6.4.2 Performance of similarity measures

The results reveal that the effect of the choice of the similarity measure is dependant upon the type of SPET image. For ^{67}Ga SPET-CT registration, the mean registration error was similar for MI, NMI and SCR irrespective of which algorithm (OSEM and FBP) was used (Figure 6.12 left column).

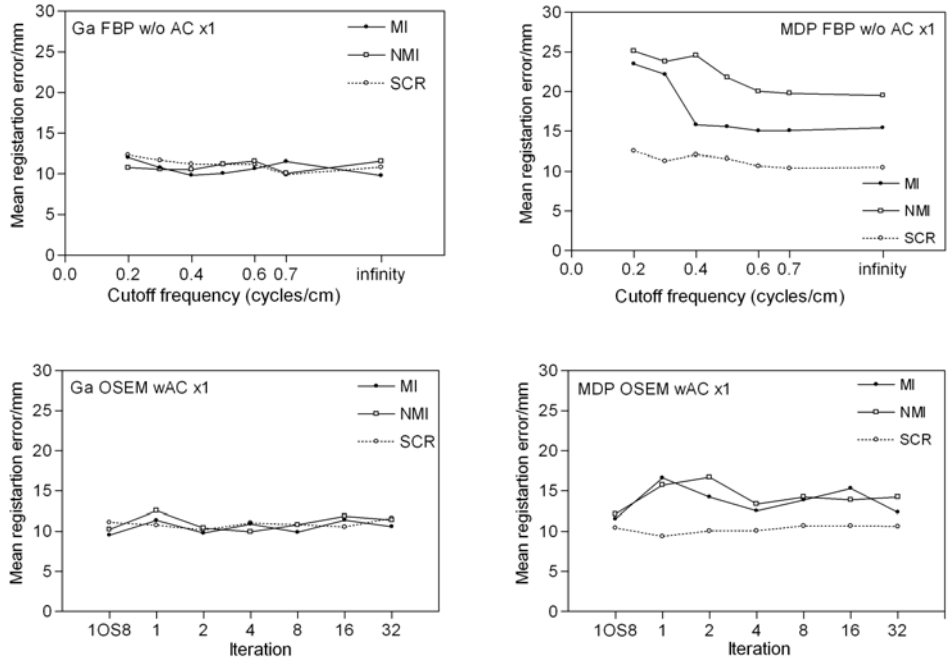


Figure 6.12: Mean registration errors (mm) for various cutoff frequencies in FBP (top row) and various iterations in OSEM (bottom row). For FBP, the abscissa gives the cutoff frequency of the Butterworth filter in cycles/cm ('infinity' refers to the unsmoothed ramp filter). For OSEM, the abscissa gives the number of OS4 iterations scaled logarithmically (1OS8 refers to a single iteration of 8 ordered subsets).

Furthermore, as was demonstrated previously for SCR (Figure 6.10), attenuation compensation in OSEM reconstruction also shows little effect, if any, on registration performance when the MI or NMI similarity measure is used. This is confirmed by the fact that there were no statistical differences at 5% level (a non-parametric one-way ANOVA test) in the mean registration error among the three similarity measures regardless of the OSEM iteration number, with and without AC. Again, a similar observation applies to ^{67}Ga FBP SPET without AC, irrespective of which cutoff frequency was used. In the case of MDP SPET-CT registration, the mean registration errors among MI, NMI and SCR were significantly different (Figure 6.12 right column), as measured by a non-parametric one-way ANOVA test (p = 0.05). This difference in the performance was also found at low projection count levels ×0.5 (not shown), with SCR consistently demonstrating a better performance than MI or NMI. For example, the

mean (±SD) registration errors (in mm) of MDP SPET-CT registration were 10.0±3.2 mm, 16.7±5.5 mm and 14.2±4.3 mm for 2OS4 with AC (×1) corresponding to SCR, NMI and MI, respectively. To further illustrate the difference of performance between MI, NMI and SCR, scatter plot of mean registration errors were given in Figure 6.13. Roughly noise-equivalent SPET images (with AC at ×1) of 2OS4 in OSEM and cutoff frequency at 0.4 cycles/cm of Butterworth filter in FBP are also shown. Again, FBP image generates streak artefacts at the border of highly active regions (spine) but OSEM image satisfactorily eliminates this artefacts.

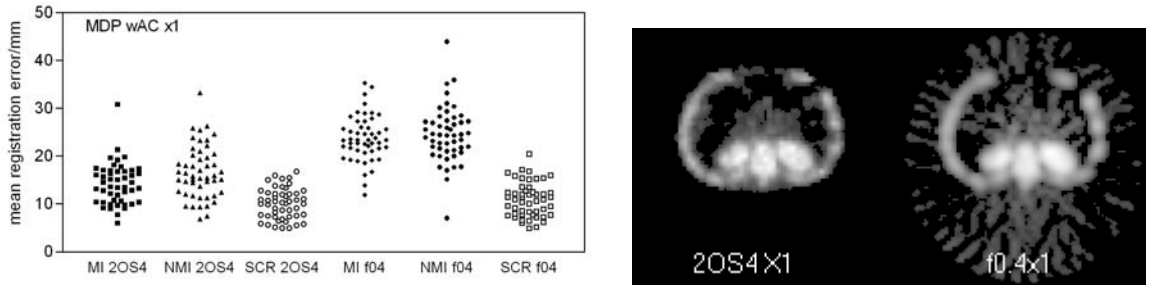


Figure 6.13: The distributions of the mean registration errors (mm) for MI, NMI and SCR for OSEM and FBP SPET-CT registration. Two noise-equivalent SPET images reconstructed with AC are also shown (OSEM on the left and FBP on the right). Each image is scaled independently.

6.4.3 Effect of background

Strictly speaking the background is always present in real clinical data, but in this experiment the background can be allowed to assume a nonzero mean count. The registration performance for projection data at high count level (×1) and low count level (×0.5), with and without background activity, is given in Figure 6.14.

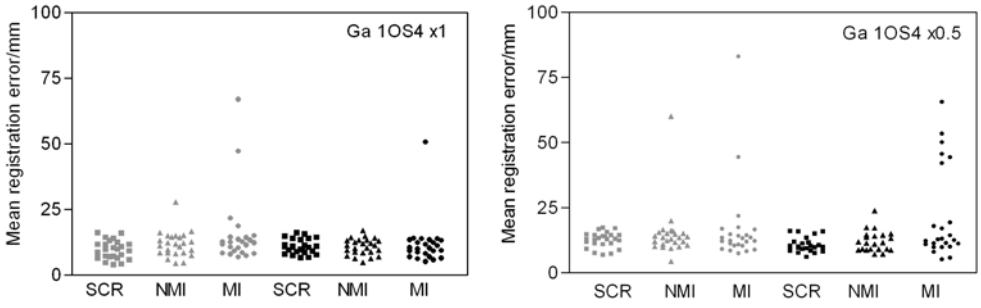


Figure 6.14: Scatter plots of the mean registration error (in mm) for OSEM ^{67}Ga SPET (1OS4) evaluated at two projection count levels. Grey dots correspond to reconstructions with nonzero background counts while the black dots represent zero background.

For SCR, the corresponding registration errors (mean ± SD) for ×1 with background was 9.7±3.6 mm and without background was 10.9±2.8mm, for ×0.5 with background

was 12.5±2.9 mm and without background was 10.6±2.8 mm. No significant difference between the mean registration errors, at the 5% level, was found between high count images (×1) with and without background activity. However, with low count images (×0.5), background activity can deteriorate registration performance of gallium-67 SPET. This is confirmed by statistical test (two-tailed unpaired t-test) which showed that the presence of the background counts significantly ($p < 0.05$) reduced the mean registration error for SCR. The performance of NMI and MI was also evaluated. Again, the presence of background activity made no significant difference to the registration error. The plot shows that, at the clinical noise level, both NMI and MI gave rise to large mean registration errors when background was included in the projection data. Removal of the background activity still produced a large mean value. To prevent possible bias of mean and standard deviation of an error distribution due to gross failure in registration (Subsection 6.2.6), mean registration error greater than 100 mm was rejected for computing summary statistics. In this case, the mean registration error (mean ± SD), with nonzero background activity, for NMI was 11.8±4.8mm and MI was 15.8±13.5 mm. The large standard deviation of 13.5 mm in MI was due to the presence of a few large misregistrations. Thus, one possible differentiating characteristic of registration performance is by counting the number of outliers. The total number of registrations in which the mean registration error was greater than the threshold (20 mm, Subsection 6.2.6) from this simulation was 1 for SCR, 6 for NMI and 15 for MI out of the total 100 independent measurements (over high count and low count image, with and without background activity). Based on this observation, SCR appeared to perform with greater robustness than NMI and MI where the failure rate for SCR was consistently lower than for NMI and MI.

6.4.4 Effect of image resolution matching

An example of a CT image blurred to match the resolution of the corresponding CT image is shown in Figure 6.15. A comparison with Figure 6.1 shows the blurring to be severe, causing small scale features to vanish. The distributions of error measured in SPET-CT registration experiments are shown in Figure 6.16.

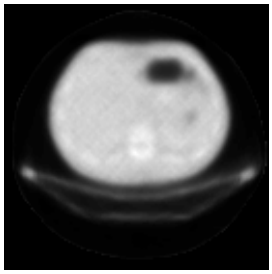


Figure 6.15: The blurred version of CT (128×128) has the same resolution as the SPET gallium-67 OS4 reconstruction

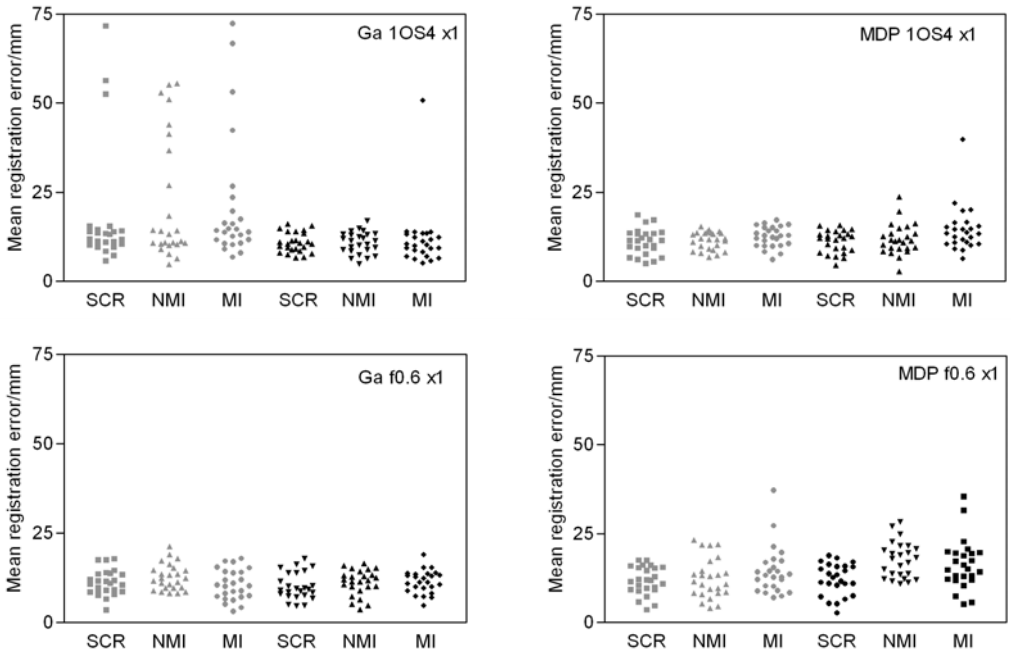


Figure 6.16: Effect of matched resolution on the performance of SPET-CT registration, for a range of similarity measures and for two radiopharmaceuticals (gallium-67 and MDP). The scatter plots in the top row correspond to OSEM, and those in the bottom row to FBP reconstructions. Default implementation was used with attenuation correction and the clinical noise level. Grey dots represent resolution matching, black dots represent unchanged resolution. For FBP, Butterworth filter of cutoff frequency at 0.6 cycle/cm was used.

The results are mixed. For reconstructions other than gallium-67 OSEM at the clinical level of planar projection count (^{67}Ga 1OS4 ×1), reducing resolution of CT to match that of SPET has no significant effect on the registration accuracy as verified by the two-tailed unpaired t-test. Taking SCR and FBP reconstruction as an example, matching resolution changed the mean (±SD mm) registration error from 10.2±3.8 to 11.5±3.5 for ^{67}Ga SPET, and from 12.3±4.5 to 11.8±3.9 for MDP SPET. However, for OSEM-reconstructed gallium-67 SPET (Figure 6.16, top left), the resolution adjustment is associated with a substantial increase in the number of outliers. The outliers (errors ≥ 100 mm were removed to prevent bias) raised the mean and standard deviation of the

mean (±SD mm) registration error from 10.9±3.1 to 22.1±17.5 (NMI), and from 11.8±8.6 to 21.3±17.9 (MI). More tellingly, the failure rate, defined as the proportion of registrations that result in the mean error of 20 mm or more, rises with the introduction of resolution matching from 0 to 12% (SCR), 0 to 32% (NMI) and 4 to 28% (MI).

6.4.5 Effect of intensity windowing of CT

In this experiment, CT pixel values outside the range [800,1250] were clipped to enhance contrast. Both the contrast-enhanced CT image and the reference image were linearly rescaled to [0,255]. Registration errors are plotted in Figure 6.17. In general, the results were diverse. When CT windowing was applied to the CT image before registering SPET images at the clinical noise level with the transformed version of CT, it introduced no significant effect on registration irrespective of the similarity measures (SCR, NMI and MI), and was confirmed by two-tailed unpaired t-test ($p<0.05$). Taking SCR and FBP reconstruction as an example, CT intensity windowing changed the mean (±SD mm) registration error from 13.2±5.4 to 10.2±3.8 for ^{67}Ga SPET, and from 11.8±5.0 to 12.3±4.5 for MDP SPET. However, for OSEM-reconstructed gallium-67 SPET at ×1 (Figure 6.17, top left), registration without CT windowing is coupled with a considerable increase in the number of outliers (error ≥ 100 mm were excluded in computing summary statistics to prevent bias), and the mean registration errors were substantially increased. For example, the mean registration errors (±SD mm) increased from 10.9±2.8 (with CT windowing) to 19.6±15.9 (without CT windowing) for SCR, 10.9±3.1 to 27.4±21.1 for NMI, and 11.8±8.6 to 19.1±15.0 for MI. The large SD was due to the presence of outliers. In this particular case, the proportion of outliers out of 25 registrations (the failure rate) changed from 0 to 28%, 0 to 36% and 4 to 28%, for the respective similarity measures; SCR, NMI and MI. That is, about a quarter to a third of registrations failed by more than 20 mm, as measured by the average displacement of the box vertices. Interestingly, CT intensity windowing shows little effect on the registration accuracy for MDP OSEM-reconstructed SPET, and this is confirmed by the statistical testing at 5% level.

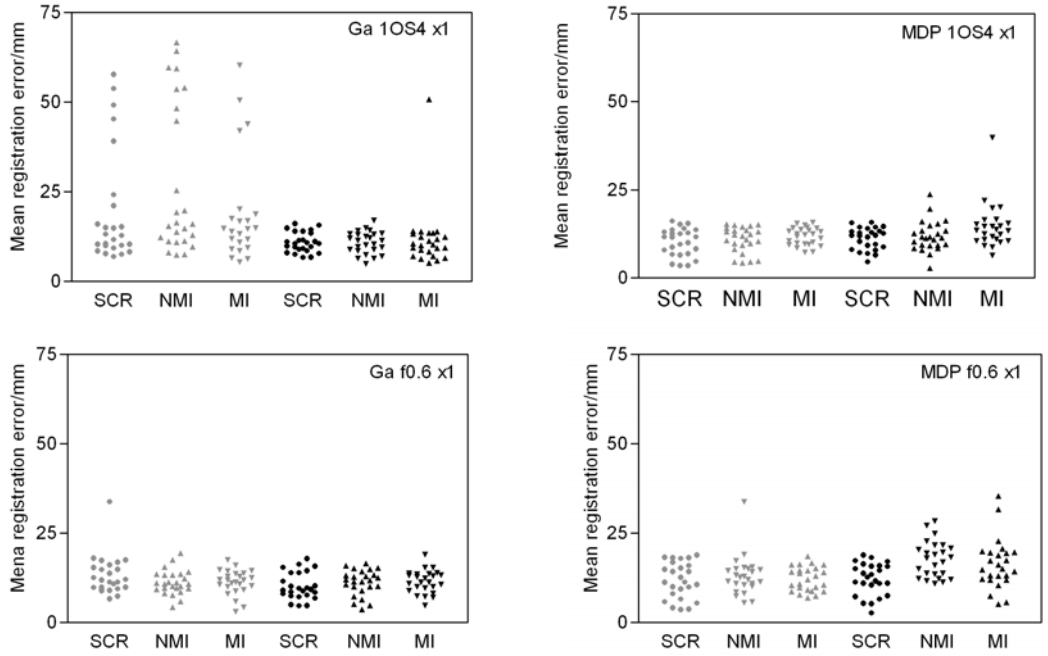


Figure 6.17: The scatter plots in the top row correspond to OSEM, and those in the bottom row to FBP reconstructions. For OSEM, 1 iteration of OS4 was used. For FBP, Butterworth filter of cutoff frequency at 0.6 cycle/cm was used. Default implementation was used with attenuation correction and the clinical noise level. Grey dots represent no CT intensity windowing, black dots represent with intensity windowing.

6.4.6 Effect of body boundary in SPET

SPET is generally associated with the poor definition of body boundary in SPET data, and the aim of this experiment is to determine whether poor boundary affects registration. Depending on the application and the type of radiopharmaceutical used, body boundary might not show up at all, or be poorly defined (Figure 6.18).

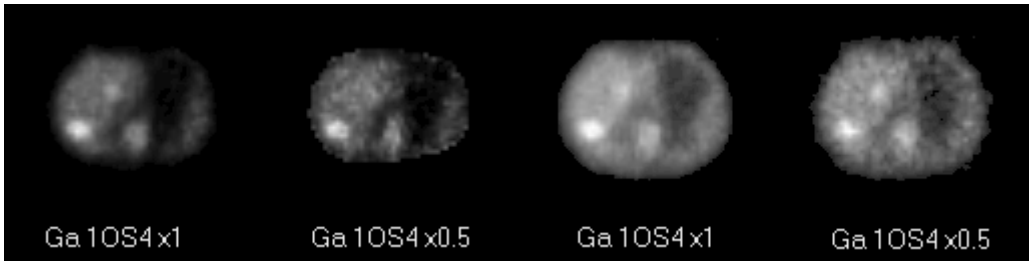


Figure 6.18: A slice of the gallium-67 SPET image (64×64) at two projection count levels obtained by 1OS4 reconstruction with AC, without body boundary (left) and with body boundary (right). Each image is scaled independently.

SPET images were assigned nonzero uptake in various tissues and organs defined by the anatomical labels of the Zubal phantom (see Appendix) to delineate body boundary. The statistics of the mean registration error in Figure 6.19 show the registration performance with and without body boundary. CT intensity windowing was used as it

has been shown to improve registration of gallium-67 SPET to CT (Section 6.4.5). The plots show little, if any, effect on the registration performance. The two-tailed unpaired t-test demonstrated no significant difference at the 5% level regardless of the clinical noise levels, the choice of the similarity measure or the planar projection count level, with the exception of NMI at the low projection count level (×0.5) in which the mean (±SD) mm registration error change from 12.1±4.0 (with body boundary) to 14.3±3.1 (without body boundary). The distributions of error for SCR and NMI were, in general, more tightly bound than for MI. The mean (±SD) mm registration error for gallium-67 SPET reconstructed with OSEM at the high projection count level (×1) deteriorated from 10.9±2.8 (with boundary) to 11.3±3.4 (without boundary), from 10.9±3.1 to 12.6±2.6, and from 11.8±8.6 to 13.0±2.5, for SCR, NMI and MI, respectively, but the changes in the mean registration error were not statistically significant.

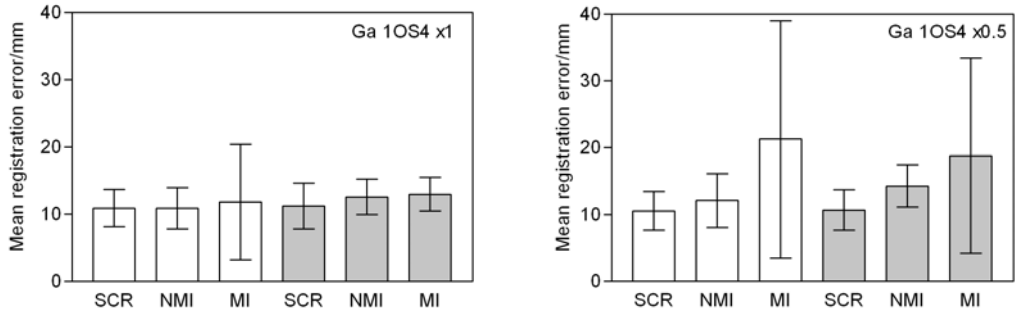


Figure 6.19: Effect of the presence (solid white) body boundary on the registration performance in gallium-67 SPET, with high (×1) and low (×0.5) projection counts. Each bar represents the mean registration error averaged over 25 SPET-CT registrations and the error bars show the standard deviation of the error distribution.

6.4.7 Effect of absence of lung in CT

The statistics of the registration error are compared in Figure 6.20 so that the effect of excluding the lungs from the CT field of view is assessed. Results show that contracting the CT field of view to exclude the lung region did have a marked effect on SPET-CT registration, particularly for the MDP SPET simulation. FBP reconstruction of MDP SPET appeared to be more vulnerable, recording the largest errors for all three similarity measures. As an illustration, lung exclusion led to the mean (±SD mm) registration error for OSEM-reconstructed gallium-67 SPET rising from 10.9±2.8 to 17.4±10.1 for SCR, from 10.9±3.1 to 14.2±4.3 for NMI, and from 11.8±8.6 to 18.6±12.4 for MI. Changes were smaller for gallium-67 SPET and no significant differences (at the 5% level) were found for FBP-reconstructed gallium-67 SPET (NMI

and MI). In OSEM-reconstructed MDP SPET, SCR was the only similarity measure that did not manifest a significant deterioration in registration performance.

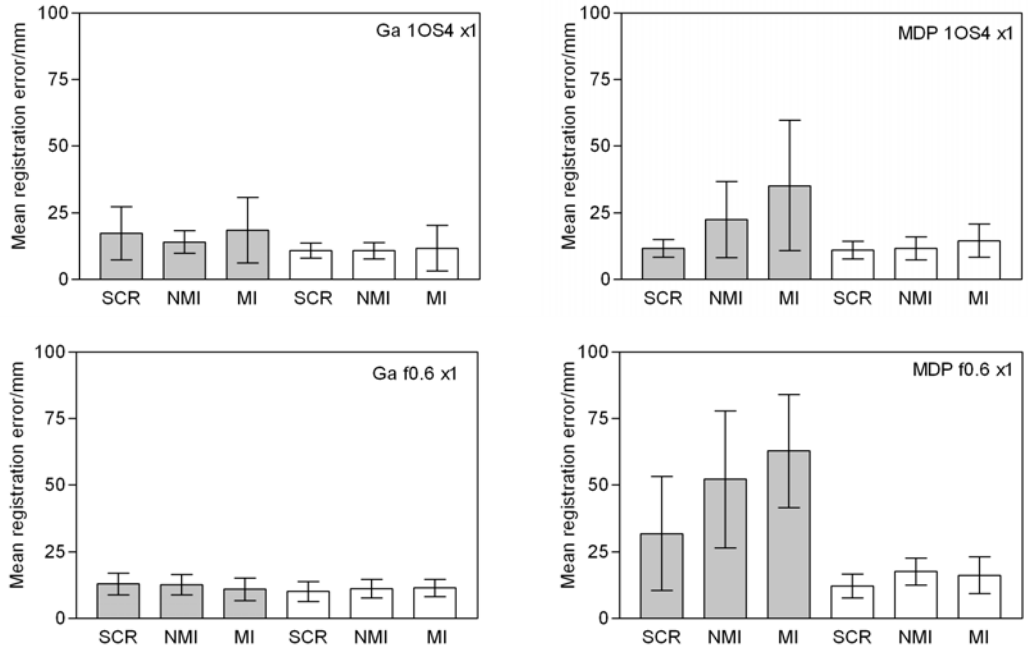


Figure 6.20: The effect of excluding the lungs (solid grey) from the CT field of view on the registration performance. Shown are the results for ^{67}Ga and MDP SPET, both reconstructed with OSEM (top) and FBP (bottom), and registered to CT using the default implementation. All graphs are shown in a common scale for easy comparison. Each data point represents an average over 25 SPET-CT registrations and the error bars represent the standard deviation of the distribution.

6.4.8 Effect of field of view

Figure 6.21 shows that registration varies with field of view (FOV) for gallium-67 and MDP, with SCR, NMI and MI, but the results are mixed. The results, with large SD for 32% FOV, is certainly poor but the distinction between the remaining means, given the high SD, is not significant statistically. These results do not appear to change significantly at larger fields of view, and is confirmed by statistical test (two-tailed unpaired t-test) that FOV has no significant effect on registration between two neighboring FOVs. Registration performance declined only when the FOV was reduced to 32% (43% at most) especially for the MDP SPET reconstructions. The number of failures, defined again as registrations with error of at least 20 mm (Subsection 6.2.6) rises sharply when FOV is at the lower end of the range examined. Results for NMI and MI, of all registrations involving MDP SPET image reconstructed with OSEM, were amalgamated, giving a total of 50 registrations. The amalgamated failure rate (the number of failed registrations out of 50 SPET-CT registrations for each FOV) was 50, 35, 12, 1 and 0 for image volume increasing from 32% to 100% progressively, but no

failure was found for SCR irrespective to the FOV. The failure rate of about 50% and higher (for the two smallest fields of view examined) cannot be tolerated in a clinical setting. Registration with gallium-67 SPET was less sensitive to the volume change than with MDP SPET, as revealed by a large difference in the failure rate: 27/75 (at 32% FOV) and 5/75 (at 43% FOV) for OSEM gallium-67 SPET, and 75 denotes registrations amalgamated from SCR, NMI and MI. For FBP MDP SPET-CT registration using SCR (not shown), the mean registration errors (±SD mm) were 14.7 ±4.3 and 11.0±3.9 for 32% and 100% FOV, respectively, and was also not significant different. The improvement for gallium-67 SPET with increasing volume at NMI and MI was less obvious. Comparing the similarity measures, SCR gave a more consistent and superior performance than NMI and MI when small FOV was used for MDP registration.

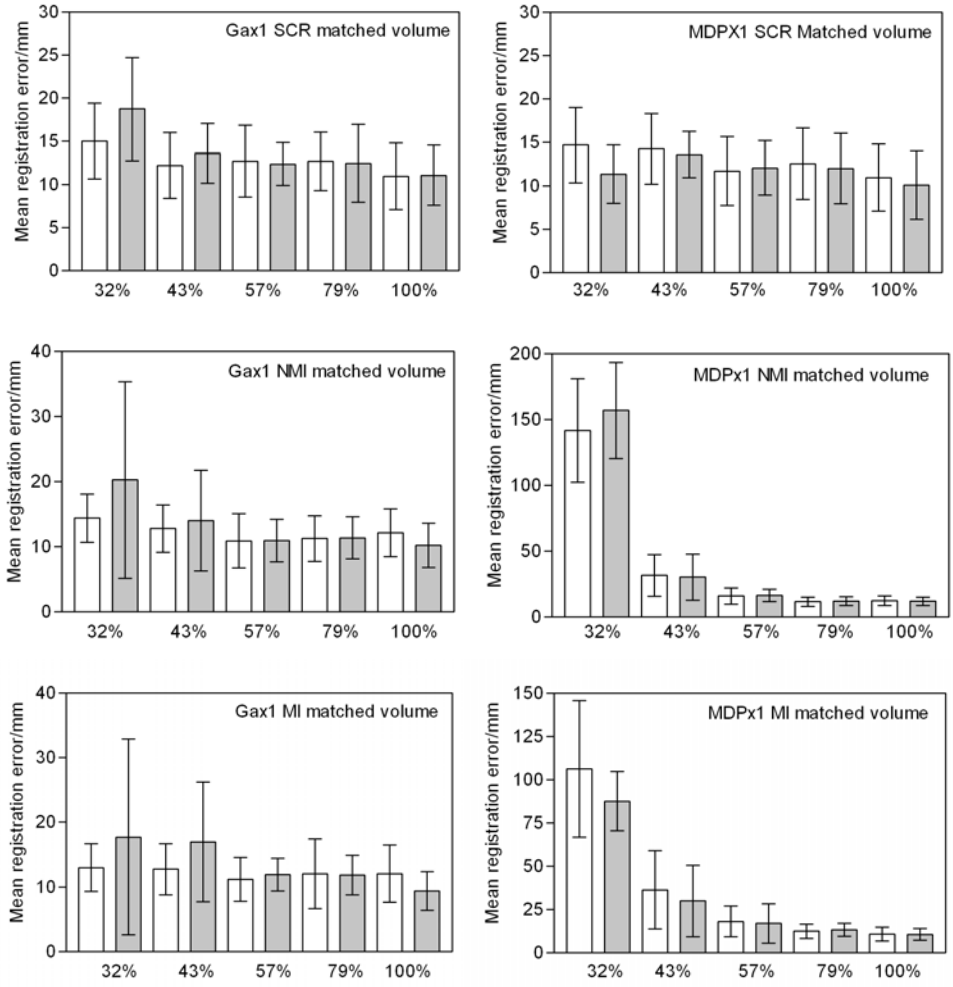


Figure 6.21: Effect of axial contraction of the field of view on registration performance. Each data point represents the mean registration error averaged over 25 registrations and the error bars give the standard variation. OSEM (solid grey) and FBP (white) reconstructions are shown side by side. Registrations based on gallium-67 SPET are shown in the left column (MDP in the right column). Each similarity measure is treated separately as indicated. Identical field of view was maintained in CT and SPET. Graphs are not drawn to a common scale.

In the above study, the field of view in SPET and CT was varied concurrently. A second study examined the effect of changing FOV in CT while maintaining the SPET field of view at 57% volume. The performance of registration is summarized in Figure 6.22.

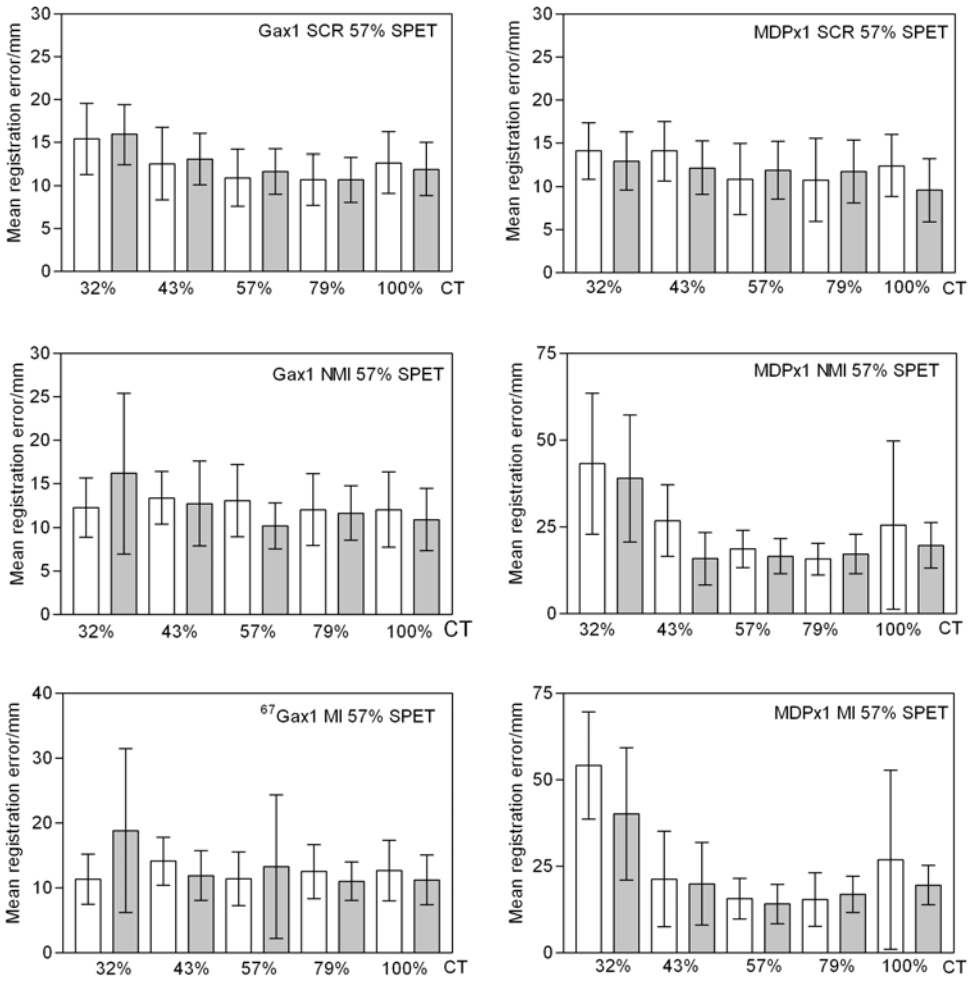


Figure 6.22 Effect on registration performance of axial contraction of the field of view in CT (x-axis) with the field of view in SPET kept at 57%. Each data point represents the mean registration error averaged over 25 registrations and the error bars give the standard variation. OSEM (solid grey) and FBP (white) reconstructions are shown side by side. Results of registrations based on gallium-67 SPET (left column) and MDP (the right column) are shown separately for each similarity measure. Graphs are not drawn to a common scale.

The results show that the overall pattern is remarkably similar to Figure 6.21 implying that the size of the FOV, as expected, is more important than the matching. However, the scale in Figure 6.22 is much smaller, implying that having one FOV at 57% reduces the number of outliers at small CT FOVs. The figure also shows the distribution of the mean registration error for registration using 32%, 43%, 57% (default implementation), 79% and 100% of the original CT FOV. The performance of registration depended on the similarity measure. The registration improved when CT FOV increased from 32%

up to 57% but its performance deteriorated with a further increase in CT FOV. The effect of CT FOV on registration was more profound in MDP than gallium-67 SPET regardless of the similarity measure. Based on the variation of the mean registration error with respect to the CT FOV, there is a hint that for CT FOV greater than 57% the mean registration error rises. For MDP SPET, the total number of failures (out of 75 registrations) was lowest at 57% FOV for MDP registration over all similarity measures. Moreover, the SD, indicated by the error bars, had less spread in case of matched 57% CT FOV. However, the change of FOV has no effect statistically on the mean registration error between 57% and FOV greater than 57%. In this case, registration accuracy appeared to be indistinguishable from 57% and beyond, although a small mean registration error was observed at about 57% CT volume compared to other CT FOV.

6.4.9 Effect of bounding box size

Given 57% field of view for the CT, it was unclear whether registration accuracy depends on the bounding box's dimension or not. This study examines two cases: the "bounding" box size changes when (1) FOV is fixed at 57%; and (2) FOV is fixed at 43%. To study the first case: SPET was registered to randomly transformed CT with 57% FOV (both CT and SPET at 57%) using default subsampling schemes, with box C_1 of size 46×31.2×43.6 cm^3 (the default size corresponding to the 57% FOV) and box C_2 with half the dimensions (23×15.6×21.8 cm^3) and concentric with C_1. The two-tailed unpaired t-test was used to check the null hypothesis of the difference in the means of error distributions corresponding to the two boxes. The results (Figure 6.23, top left) show that the box size had, unexpectedly, no significant effect on the registration accuracy at 5% level. In particular, the mean (±SD mm) registration error was 10.9±2.8 (C_1) and 12.1±2.7 (C_2) for OSEM gallium-67 and 11.3±3.1 (C_1) and 10.0±2.9 (C_2) for OSEM MDP SPET (not shown). To further analyse the residual registration error, mean registration error $\bar{\varepsilon}$ was decomposed into mean translational error $\bar{\varepsilon}_t$ and mean rotational error $\bar{\varepsilon}_R$ (Subsection 6.2.6). For example, mean translational errors (in mm) for OSEM gallium-67 SPET-CT registration were 10.2 for C_1 and 12.0 for C_2, nearly identical to their corresponding mean registration errors: 10.9 and 12.1. Similar observations were found for MDP SPET-CT registration. These results indicate the residual error was dominated by the residual translational error. An example of

registration performance for each of the six transformation parameters is shown for C_1 and C_2 boxes in Figure 6.23 (bottom).

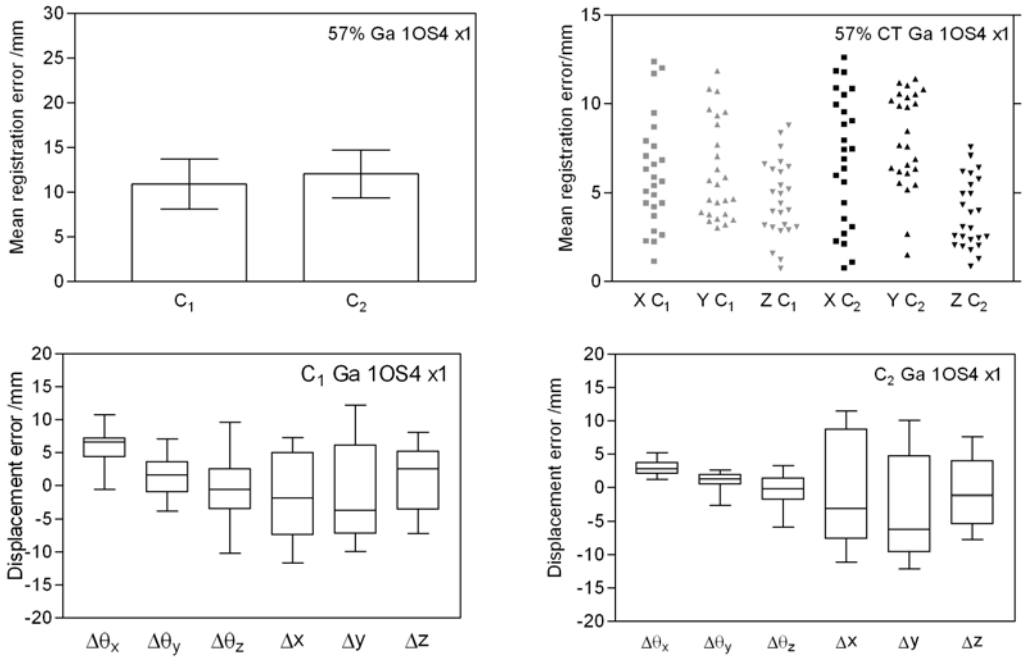


Figure 6.23: (Top left) Mean registration (±SD) error in mm measured with the C_1 and C_2 boxes. Both CT and SPET had the same 57% FOV. Default sampling scheme was used. (Top right) Vertex displacement error averaged over the eight vertices of the box along x, y and z direction. (Bottom row) displacement error distribution for registration of gallium-67 SPET (×1 wAC) to 57% CT volume, with C_1 (default box size) and C_2 (half of the default box size). In the box-whiskers plot, the central horizontal line indicates the median, the top/bottom boundary of the box give the 25 and 75 percentile, and the extreme values give the highest and lowest error value in the distribution.

In order to compare translational and rotational components of the registration error, the rotational error was given as the vertex displacement. For example, in a 2D plane, rotation about the centre of the box by displaces each vertex of the box by $2l\sin\theta/2$, where l is the distance from the box centre to its vertex. It was surprising that, unlike any other parameter, $\Delta\theta_x$ was not distributed more or less symmetrically about zero; instead, it was strongly biased towards positive values and there were very few instances of negative values. The graphs also showed that rotational error was more tightly bound than the translational error. Of the translational errors, z-axis errors were smallest. The distribution of the translational components of the registration error is shown in Figure 6.23 (top right). Note that C_1 and C_2 displayed similar patterns.

In the second case, the FOV of both CT and SPET was reduced to 43% (Figure 6.24). The mean registration errors measured with the C_1 and C_2 boxes were significantly (P=0.01, two-tailed unpaired t-test) different. In particular, the larger box showed a

larger mean rotational error (Subsection 6.2.6). For example, the mean rotational error (in mm) was 6.0 for C_1 (large box) but it was 1.8 for C_2 (small box) when 43% CT volume was used for registration. For 57% CT volume, the mean rotational error was 0.7 mm for C_1 and 0.1 mm for C_2. Again, the registration error in mm along x, y and z-axis were given in Figure 6.24 (right), where the residual error in each component had no different statistically at 5% level.

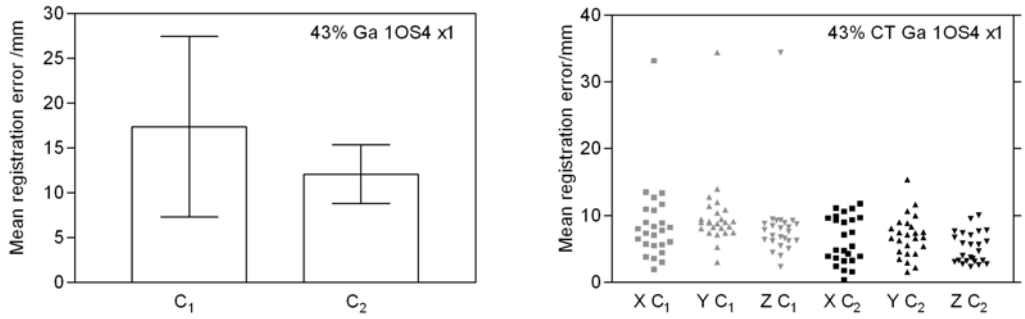


Figure 6.24: (Left) Mean registration (±SD) error in mm measured with the C_1 and the C_2 box. Both CT and SPET had the same FOV (43%). Default sampling scheme was used. (Right) Translational component of the vertex displacement error along x, y and z direction as measured with the C_1 and C_2 box.

6.5 Discussion

In the thoracic-abdominal SPET, methylenediphosponate (MDP) and gallium-67 scans were identified for investigation because they are in common clinical use. Gallium-67 citrate is used in the diagnosis of lymphoma and inflammation. MDP labeled with technetium-99m provides images with a high bone to soft tissue contrast for the diagnosis and assessment of bone disease. The MDP SPET is often referred to as the bone scan.

6.5.1 Reconstruction algorithm

Given the utilization of two main reconstruction algorithms (FBP and MLEM) in clinical SPET, the question of whether the choice of the algorithm affects the registration process warrants an investigation. However, a comparison is not straightforward as the algorithms produce images of different quality, and each requires the setting of different parameter values. In OSEM, once the number of subsets is chosen, the number of iterations entirely controls the final resolution, contrast and noise of the reconstructed images [348]. In FBP, for a given choice of the low-pass filter, its cutoff frequency controls the quality of the reconstructed image. The reconstructed image noise is likely to be the limiting image quality factor in image registration

suggesting that reconstructions be obtained at the same noise level to enable a fair comparison of reconstruction algorithms. Because it is not straightforward to predict a "noise level" from the setting of these two parameters (number of OSEM iterations and FBP low-pass cutoff frequency), a different strategy was adopted. Reconstructions were obtained for a range of parameter values, and noise level was measured in the resulting image (Subsection 6.3.1). From this set, pairs of images were selected, which had approximately the same noise level. Thus the comparison between the registrations involving FBP and OSEM reconstructions was based on noise-equivalent images. Mean registration errors were measured for each reconstructed SPET as a function of reconstructed noise, with and without AC. The results show that the choice of the reconstruction algorithm did have a significant effect ($p<0.05$, two-tailed unpaired t-test) on registration but only if FBP reconstruction incorporated attenuation correction. However, the assertion could only be tested over a small range of noise levels due to the small number of noise-equivalent reconstruction pairs available. This, to some extent, supports the assertion that OSEM produced noisier images and led to more accurate registration.

It is also worth noting that although AC is essential for removing attenuation artefacts from reconstruction, inclusion of AC in OSEM reconstruction had little effect, if any, on the SPET-CT registration accuracy (Subsection 6.4.1). By contrast, as pointed out above, inclusion of AC in FBP reconstruction (of ^{67}Ga or MDP SPET) had a significant deleterious effect on registration accuracy. This could be related to the AC algorithm for FBP reconstruction. In particular, in Chang's algorithm [314] used here, a multiplicative attenuation correction is applied. Without AC in FBP, noise is relatively constant throughout the image. With AC, noise varies across the image perturbing the joint intensity histogram. As a result, similarity measure (e.g. NMI) could show a shift in optimum, as illustrated in Figure 6.25, for a one-dimensional translation for FBP gallium-67 SPET-CT registration. The position of zero translation corresponds to the true alignment. The diagram reveals that, with AC, the correct registration position is shifted. This helps to explain why the registration for gallium-67 FBP-reconstructed SPET was more accurate without AC than with it (Subsection 6.4.1). In registration of bone SPET (MDP) to CT, OSEM consistently outperformed FBP-reconstructed bone SPET using MI and NMI (Subsection 6.4.2). A similar observation that the OSEM reconstruction method gave better results to those of the FBP method in bone SPET was

also reported in [340]. This is due to the fact that the FBP reconstruction has some disadvantage in bone SPET. The most important shortcomings are streak artefacts and negative values in the border of highly active bone regions. In contrast, OSEM is able to eliminate, at least partially, almost all artefacts observed with FBP bone SPET (Figure 6.13).

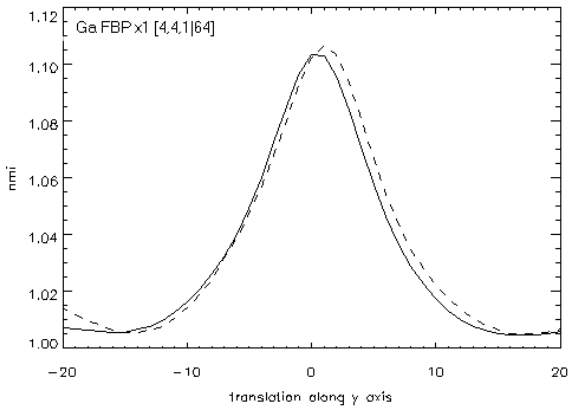


Figure 6.25: NMI similarity measure as a function of 1D translation for SPET-CT registration, where gallium-67 SPET is reconstructed using FBP without (solid line) and with (dashed line) attenuation correction.

6.5.2 Similarity measures

The performance of three similarity measures: SCR, NMI and MI, was compared in a number of experiments. Generally, for gallium-67 SPET, the SPET-CT registration performance is similar for all three measures and is not affected by the choice of the reconstruction algorithm. However, in registration of bone SPET (MDP) to CT, SCR consistently outperformed MI and NMI (Subsection 6.4.2). One-dimensional results (Figure 6.26) also suggest that SCR is better defined than either NMI or MI, containing fewer spurious maxima, and is more likely to converge from an initial estimate to the maximum, yielding the correct transformation parameters.

In considering the success or failure of registration, it is helpful to consider two cases: fine and gross residual misregistrations. Fine misregistration arises when the registration algorithm brings the image close to the correct alignment but residual error remains might due to the limitations of image quality (e.g. resolution, noise, artefacts) or the registration algorithm (e.g. similarity measures, sampling density). Fine residual errors can be meaningfully captured by statistical moments, such as the mean and standard deviation of the distribution. The gross misregistration, on the other hand, represents the

cases, where the registration algorithm terminates at a local optimum, which is unrelated to the correct alignment. Including such cases in the statistical analysis will strongly bias the statistics. Even the median error may be affected if gross residual errors dominate a set of registrations. Gross registration is more meaningfully measured by the rate at which it occurs. This failure rate is indicative of the reliability of the registration method and would be an important consideration in deciding whether to adopt the method in clinical practice. One difficulty with this classification is that it relies on the threshold value of error. It is likely that a useful threshold could be constructed for each clinical application, based on the clinical requirements. If the clinical rationale for registration is to distinguish between closely spaced features (e.g. to determine whether a tumour found in SPET is located in the rib bone or in the lung), the threshold will need to be set low. A contrary example may arise where the issue of interest is whether the uptake in a large organ is abnormally high. In such cases, the threshold could be raised without unduly compromising the diagnostic accuracy. In the simulation studies described here, the threshold was set somewhat arbitrarily to 20 mm for the sake of comparison between similarity measures. The failure rate, defined as the number of registrations with the mean error of 20 mm or more, was used as a measure of robustness.

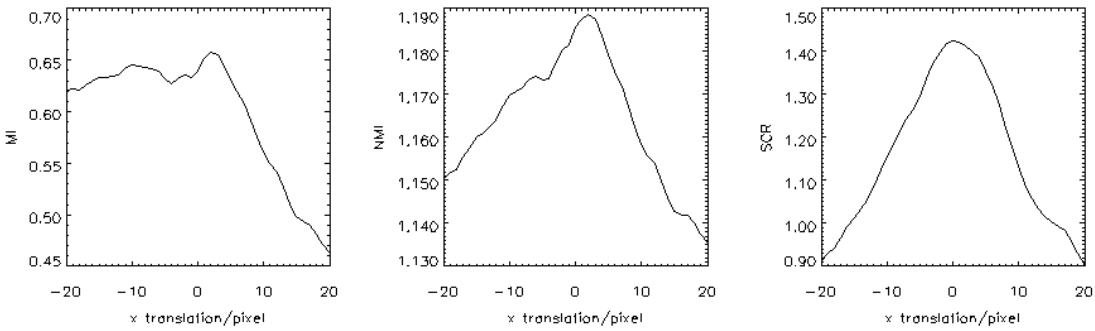


Figure 6.26: A comparison of (from left) MI, NMI and SCR as functions of *x*-axis translation for the FBP-reconstructed MDP SPET (Figure 6.13 rightmost image) with respect to the CT (Figure 6.1 left image). Full sampling was used. Correctly registered position is zero.

6.5.3 Image quality

One difficulty for SPET-CT registration is the difference in edge quality at the body boundary in CT and SPET. In CT, the body boundary is characterised by a sharp transition. By contrast, SPET exhibits a gradual change, if any (where contrast is zero). By reducing CT resolution and matching to the image resolution of SPET image, registration accuracy may be expected to improve. Results of measurements show that a

reduction of CT resolution to match that of SPET had little effect on registration accuracy. In the default setting, the raw SPET images were converted to a byte array by linear mapping the voxel values to 256 grey levels without clipping. The only effect one might expect on the joint histogram would be due to quantization errors (±binsize/2), which may eliminate small contrast features. The raw CT was also converted to a byte array using linear intensity rescaling in two cases: with and without clipping (CT intensity windowing). Because CT intensity re-mapping from intensity interval [800,1250] to [0,255] is a linear operation, it has little effect on the histogram but clipping of intensities outside the interval may have a significant effect. One might expect that, due to the large number of pixels that appear with value 0 and 255 in the transformed image, the joint histogram may have large values at the first and last column or row. Of course, the co-occurrences that corresponded to pixel values outside the interval [800,1250] are now compressed into these two columns/rows. Intensity windowing will enhance visual contrast for tissues in the range [800,1250]. Aside from quantization errors, one may expect the improved visual contrast to have significant effect on the joint histogram by stretching intensity values along the histogram axes. In fact, since the tissues that now have a uniform value of 0 or 255 carry zero (Shannon's) information, the effect of clipping the image on the histogram would be equivalent to excising the clipped parts of the image outside intensity interval [800,1250]. With a correct choice of pixel intensity interval, the destruction of contrast for tissues with intensity values outside this interval is unlikely to be significant provided that no prominent image feature is truncated or lost. Furthermore, tissues that have intensity values within the interval [800,1250] will show up visually and increases the Shannon information content for registration by spreading intensity values over [0,255] interval. This helps to explain, based on the results in Subsection 6.4.5, a significant difference was found for OSEM gallium-67 SPET-CT registration irrespective of the similarity measure. The results indicate that, because CT intensity windowing improves contrast in soft tissue, it can improve registration if activity uptake is found in such tissues. However, CT intensity windowing made no significant difference for registration with OSEM MDP SPET. An MDP image contains little soft tissue activity. It is characterised by clearly delineated landmarks (i.e. bone structure) with good uptake in the bone, giving a good correspondence with bone structures in the CT image. Even without intensity windowing, bones appear as prominent features in CT (Figure 6.1) that ought to provide sufficient information to drive the registration algorithm to the optimum.

What's more, with the absence of soft tissue uptake to define body boundary, CT intensity windowing plays a critical role for a one-dimensional translation for OSEM gallium-67 SPET-CT registration to give a correct registration position, otherwise the local maxima is shifted producing spurious maxima (Figure 6.27). In this case, CT windowing helps constrain the similarity measure.

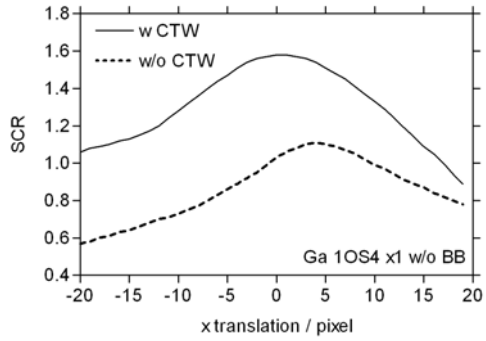


Figure 6.27: Effect of intensity windowing in CT (CTW) on SCR as function of *x*-axis translation for the gallium-67 OSEM-reconstructed SPET (without body boundary, Figure 6.18 leftmost) with respect to the corresponding CT image (Figure 6.1). Full sampling was used. Zero translation denotes correct registration.

A few studies included body boundary information for registration but no comparison had been made to study the case of registration without body boundaries. For example, Dey *et al.* [260] and Suga *et al.* [347] suggested the use of body boundaries to improve registration between functional images and anatomical images. Dey *et al.* used transmission scans while Suga *et al.* used Compton scatter to define the body boundary. Backscatter sources were also utilized [280] to obtain body contour for registration. Based on the results of registration comparison with and without body boundary made in Subsection 6.4.6, body boundary, caused by uptake in skeletal muscle and superficial tissues, in SPET image did not significantly reduce the mean registration error with CT windowing. This may due to the fact that organs or tissues deep inside the body could provide sufficient information content for registration, without the presence of the boundary information.

6.5.4 Field of view

The effect of changing the field of view in the above studies was measured using the mean residual displacement of the vertices of the bounding box. However, the box was fashioned for the 57% field of view (FOV) and, in fact, was no longer bounding when FOV was increased to 79% and 100%. In those cases, it could no longer be said that the

error provided an upper bound for the displacement error of all image voxels. On the other hand, when the volume was reduced below 57%, the calculated error became also less indicative of the voxel displacement error.

In a clinical situation, the fields of view in CT and SPET rarely coincide. The results reported in Subsection 6.4.8 shed light on this problem. Two experiments were conducted: (1) both fields of view were concurrently varied; (2) the field of view in one modality was fixed while it varied in the other modality. Since changes in the results do not appear to be significant at larger FOV, there is little, if any, correlation between mean registration error and "increasing volume". This observation can be explained by the fact that image volume affects the joint histogram, although it may not be the sole factor, and which tissues are involved do matter. A larger FOV generally contains more information for image alignment. By contrast, a smaller FOV produced larger mean registration errors and greater standard deviations of the error distribution. This implies that one should avoid too small FOVs. Thus, registration performance declined when the FOV was reduced to 32% (43% at most) especially for the MDP SPET reconstructions. When the FOV is continually reduced, the registration would be expected to eventually fail (e.g. histogram-based methods will run out of data). What matters here is whether the failure will occur with clinically realistic FOV and, if so, which methods may be more vulnerable. In general, the similarity measures (SCR, NMI and MI) have a very similar performance except for MDP (bone) SPET of small FOV. The results indicated that, with MDP SPET, SCR was more robust (i.e. smaller failure rate) than NMI and MI when small FOV were concurrently varied. The results of these measurements also demonstrate that MDP was more sensitive to the image volume change than gallium-67 SPET, when NMI and MI were used for registration. Interestingly, MDP was less robust in the lung exclusion experiments implying that this owes much, if not all, to the reduction of FOV rather than exclusion of a particular region.

It is generally difficult to match the field of view in SPET and CT. The second investigation simulated the case of mismatched fields of view. A 57% SPET volume was registered with a range of CT volumes, progressively increasing in the axial direction. The proposition was that irrelevant information in unmatched parts of the image volume could affect registration. In particular, tissue types in CT with no

corresponding SPET intensity could generate additional local maxima for the similarity measure. No significant effect was found for gallium-67 SPET and for most of MDP SPET experiments. However, when MDP SPET, with FOV of 57%, was matched to CT (32% FOV), the registration error doubled, at least when using NMI and MI. The relative susceptibility of bone SPET could be related to the fact that it is more like CT in contrast and is therefore more likely to develop strong spurious, as well as genuine, matches. The results also suggest that voxels from the same anatomical structures, to some extent, were necessary to construct a joint intensity histogram of sufficient information contents to achieve an optimum registration. The presence of unmatched tissues or organs, such as lung, in the two images may present a problem for fully automatic registration.

Registration error used in this chapter was defined as the mean residual displacement of vertices of a hypothetical box drawn around the body (Subsection 6.2.6). A smaller registration error is expected from a smaller box since the rotational error depends on the distance of the vertices from the box center. On the other hand, the registration error due to the residual error in translation is independent of the size of the box. The results of Subsection 6.4.9 reveal that the mean registration error was very similar to the mean of the translational error regardless of the size of the box used to compute the error, so long as the CT volume was relatively large (57%). Since the mean registration error did not vary with the size of the box, rotation could have only a limited effect on the registration accuracy and the error was predominantly due to the translational misregistration (Figure 6.23 bottom). Alternatively, the method of estimating the registration accuracy might not be sensitive to the box size. However, when a smaller CT field of view (43%) was used for registration, the size of the box did affect the registration error. This may mean that, for a reduced volume of interest, the registration error was more likely to be affected by the misregistration in rotation. Results show that mean registration error was reduced when the hypothetical box dimensions were halved even though the translational residual error was unchanged. Thus, for a large image volume, the mean registration error is dominated by the translational residual error whereas for a small image volume, the rotational residual error becomes more significant. One possible explanation relates to the method of varying the FOV. Larger volumes were more axially extended and therefore tend to reduce rotation about x and y-axes. For the rotation about the body axis, the measuring box is also extended axially

so errors due to axial rotation would also be smaller. And clearly, changing the box size will give different results for mean registration error for small FOV. No attempt has been made to derive a systematic correction of the mean registration error as a result of changing the FOV. For experiments involving FOVs, a rough correction for volume size could be made by assuming that the average displacement is all due to rotation about box centre. This may probably give a better estimate of the error statistics.

6.5.5 Shortcomings and further analysis

Several aspects of practical significance need to be considered in the registration of thoracic-abdominal SPET and CT images. In the absence of effective immobilization, the patient's body tends to show a relative rotation about the axial direction between MR/CT and PET images [349]. To reduce the differences in pose, Forster *et al.* [76] used a patient specific cushion with fiducial markers fixed on it to replicate positioning in SPET and CT. Motion of organs induced by respiration is also an issue for image registration [97]. The top of liver has been reported to move as much as several centimeters during respiration [350] but it is unknown how much movement occurs in other parts of the liver and in other tissues. In SPET, such movement can induce motion artefacts. The effect of motion artefacts on registration requires further study. The effects of organ deformation and of the relative movement among abdominal organs have not been fully evaluated. A quantification of these movements could improve the SPET phantom simulation. To account for the local displacement of internal organs between scans carried out using different positioning protocols, rigid-body registration is not adequate. An affine transformation can be applied as a first approximation, then global warping is used for an intermediate correction and finally some local non-linear warping model is probably required. Because a large image volume is less sensitive to local misalignment, another possible approach is to divide the image into sub-regions small enough to detect local transformations. The results of this study highlights that the presence of body boundary, a large field of view and CT intensity windowing, all have a positive effect on registration. These results should carry over to local registration. Thus the presence of boundaries, including boundaries of internal organs, will probably improve local registration of sub-volumes that contain such boundaries. Image segmentation, although it can be as difficult as the registration problem itself, is one of the possible methods to improve registration. Further investigation is needed to

investigate the optimum size of the sub-volume for the purpose of registration and the factors that control it.

6.5.6 Magnitude of registration error

In this study, the mean registration errors were found, on average to be 11 mm, larger than the largest voxel dimension (8 mm for SPET), contrary to expectation. A large error may be due to the dimension of the box used in estimating the error or to interpolation errors and resolution effects, and in a number of cases, it is due to gross registration failure. It is worth noting that where the box represents a bound on the image volume, the maximum displacement of the vertices gives the upper bound on the displacement of all voxels, leading to an overestimation of the typical registration error. The absolute magnitude of error may not be a problem when the purpose of investigation is to examine relative effects of various factors. However, if the results are compared with those computed by others, using different methods, this approach might not give a fair assessment. For the purpose of measuring the highest attainable accuracy of registration for gallium-67 and MDP SPET, further experiments were conducted.

The registration of the CT with itself probably represents the most favourable case for the joint intensity based approach, as the intensity values are perfectly dependent between the two images. Any deviation from the ideal case, as a result of smoothing, image degradation, sub-sampling and interpolation, could deteriorate the registration accuracy. In order to avoid confounding effects of differences in noise, resolution and intensity, the CT image was registered with itself using the default implementation (Subsection 6.2.5). In addition, to investigate the effect of resolution, the CT image was registered to three different versions of itself: (1) a blurred version of CT (128×128×243) using a boxcar filter (5×5×5), (2) a trilinear interpolated version of CT (64×64×121), and (3) a trilinear interpolated CT (64×64×121) followed by a boxcar smoothing (5×5×5). The trilinear interpolation and boxcar filter helps to reduce the resolution of the original CT image. The error distribution of the CT-CT registration of the four cases is given in Figure 6.28, for each of the three similarity measures.

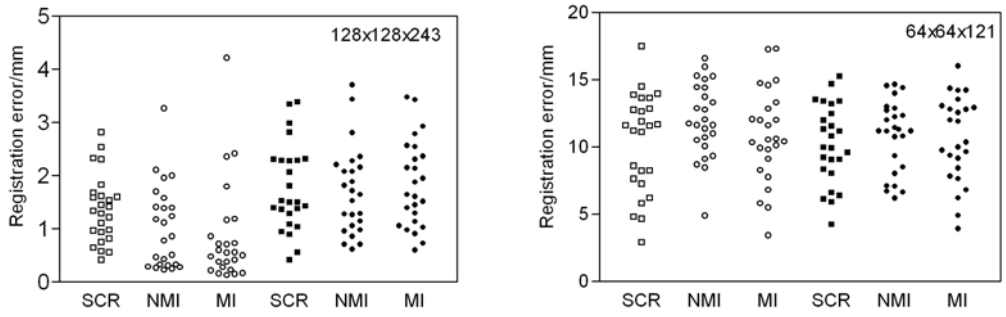


Figure 6.28: Scatter plot of registration error when CT (128×128×243) was registered to itself (grey dots) and a blurred version of itself (black dots) for the indicated similarity measures is shown on the left. Results of the registration when the CT registered with a resized version CT (64×64×121) and the results of registration without (grey dots) and with (black dots) blurring the image are shown on the right.

Using the 128×128×243 matrix, the mean registration error was found to be less than the pixel size (4mm), irrespective of blurring, and for every similarity measure. For example, with SCR, the mean (±SD) error in mm without blurring was 1.4±0.6 compared to 1.8±0.8 with blurring. Likewise, for NMI, it was 1.0±0.8 to 1.7±0.8, and for MI, 0.9±0.9 to 1.8±0.8. Except for SCR, registration accuracy between CT with and without blurring was significant at 5% level (two tailed unpaired t-test). It signifies that registration could be worse by the reduction of image resolution. This is further supported by the registration results obtained by registering the original CT with a reduced resolution version of CT (64×64×121) using SCR, NMI and MI, with and without blurring. The reduced version of CT was obtained by trilinear interpolation of the original CT. In this case, the mean (±SD mm) error, without blurring, was 10.4±3.7, 12.1±2.8 and 10.9±3.5 for SCR, NMI and MI, respectively. Again, with blurring, the mean (±SD mm) error was 10.2±3.0, 10.9±2.7 and 10.6±3.2 for SCR, NMI and MI, respectively. No statistically difference was found between theses two cases, with and without blurring. When comparing the results of the registration using CT sized 128×128×243 to those reduced resolution of CT using 64×64×121, a significant (p<0.0001, two-tailed unpaired t-test) difference was found for each similarity measure. It suggests that reducing image resolution has a profound effect on registration accuracy. It also reveals that attainable accuracy for the SPET-CT registration accuracy from experiments described in this chapter is around 11 mm. This helps to explain mean registration error for gallium-67 and MDP SPET-CT registration obtained in this chapter is close to the limit of accuracy attainable. The same accuracy (10.6 mm) was

also obtained by Skalski *et al.* [351] using mutual information-based similarity measure for registration of CT and PET in the thorax.

6.6 Clinical example

The following three cases demonstrate automatic volumetric registrations in the thoraco-abdominal region using SPET/CT and PET/CT data sets [54,352]. In the first case, CT data, acquired as 512×512 image matrices of 0.68×0.68 mm^2 pixels, representing contiguous 10 mm thick slices, were first downsampled to 256×256 matrices of 1.3×1.3×10 mm^3 voxels to reduce the computational burden. The SPET scan with iodine-123 metaiodobenzylguanidine (^{123}I-MIBG) was acquired using 128×128 matrices (4.7×4.7 mm^2) with 38 slices of thickness 9.3 mm. Registration was performed from SPET (floating image) to CT (target image) using rigid body transformation. Figure 6.29 shows a transverse slice of CT, SPET and overlays of the CT on the registered SPET image. The area with increased receptor binding in the liver is exactly superimposed on the faintly visible lesion of a carcinoid tumuor on CT. At a later time, SPET showed an accumulation of receptor binding in the liver and activity distribution of ^{123}I-MIBG in other parts of the body (e.g. the kidney). In addition to the lesion, the margin of the liver is well registered.

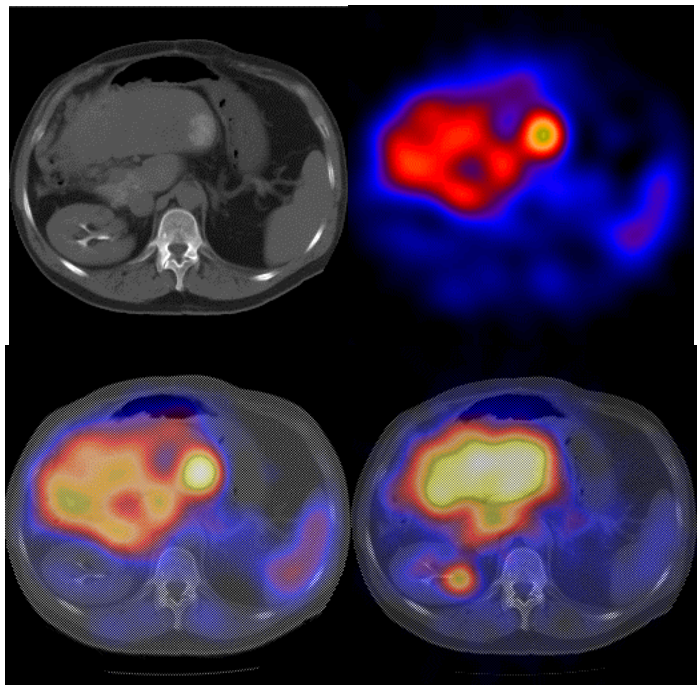


Figure 6.29: Transaxial images, following registration, of a patient with liver metastases of a carcinoid tumor. (Left panel) one metastasis is faintly visible as a brighter region in the CT image (top left) and as a hot spot in the corresponding MIBG SPET image (top right). Fusion images obtained after 16 h (bottom left) and 24h later (bottom right).

The second case demonstrates registration of 18-fluorodeoxyglucose (FDG) PET and CT. The PET scan traces glucose metabolism *in vivo*. This patient was diagnosed with liver metastases (hepatoblastoma) and confirmed by Gallium abnormal uptake. The patient received chemotherapy and was subsequently operated on the right liver (hemi-hepatectomy). However, blood chemistry demonstrated a slowly rising serum liver enzyme concentration implying the possible of recurring of the liver lesion. Further diagnosis using a gallium-67 scan was negative, but ^{18}FDG scan was positive. The critical question was whether the lesion found in PET was due to post-surgical scarring or recurrence of hepatoblastoma. CT shows both of metastases or scarring while FDG only indicates metastases. In terms of image registration, the question being posed was whether the CT and FDG lesions coincided. A registration of fluorine 18-fluorodeoxyglucose (FDG) with CT was performed showing recurrence of the metastases. The lesions detected by FDG appeared exactly superimposed on the respective CT lesions by visual inspection, which was later confirmed hepatoblastoma by biopsy. Figure 6.30 illustrates the fused datasets of PET-CT scans.

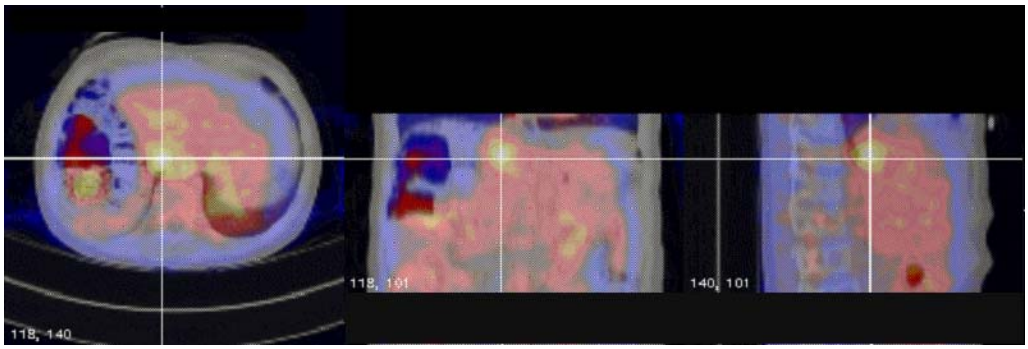


Figure 6.30: The area of recurrence of hepatoblastoma was unremarkable in CT. A study of fluorine 18-fluorodeoxyglucose (^{18}FDG) was performed showing focal localisation of the disease. Fusion of FDG and CT is shown in three orthogonal sections (left to right: transverse, coronal and sagittal). Increased uptake of ^{18}FDG in the receptor area indicated the corresponding location in the CT.

The third case demonstrates registration of gallium-67 SPET and CT in the chest that provided information with significant clinical implications. This patient had a history of complex congenital heart disease, requiring several cardiac surgical procedures to insert a bioprosthetic pulmonic valve and pulmonary outflow trunk. Seven months earlier the patient had a patch repair of the ventricular septum requiring a surgical procedures by formation of a right ventricle to pulmonary artery conduit. Later, the patient presented with high fever and blood-stained sputum. Bacteraemia (Staphylococcus haemolyticus) was cultured from blood, and was suspected of causing prosthetic infection. A gallium-67 scan was requested. After intravenous injection of ^{67}Ga-citrate, whole body planar

and chest SPET images (64×64 matrices) were acquired 48h later after injection. Planar images showed a focus of mild gallium-67 localisation in the right mid thorax near the mediastinum. The distribution of gallium-67 elsewhere in the body appeared unremarkable. There was insufficient structural detail to establish confidently whether the hot spot was due to the intrathoracic lesion in bronchopulmonary tissue or a pulmonary vessel. The critical distinction to be made was between pneumonia and prosthetic vascular graft infection. When the registered images were reviewed (Figure 6.31), it was clear that the gallium-67 sensitive lesion was outside the lung tissue and within the prosthetic vascular conduit. An inflammatory focus labeled by gallium-67 was seen in the thorax and the SPET-CT registered image was conclusive in localising the lesion to the vascular prosthesis. This was confirmed by further imaging.

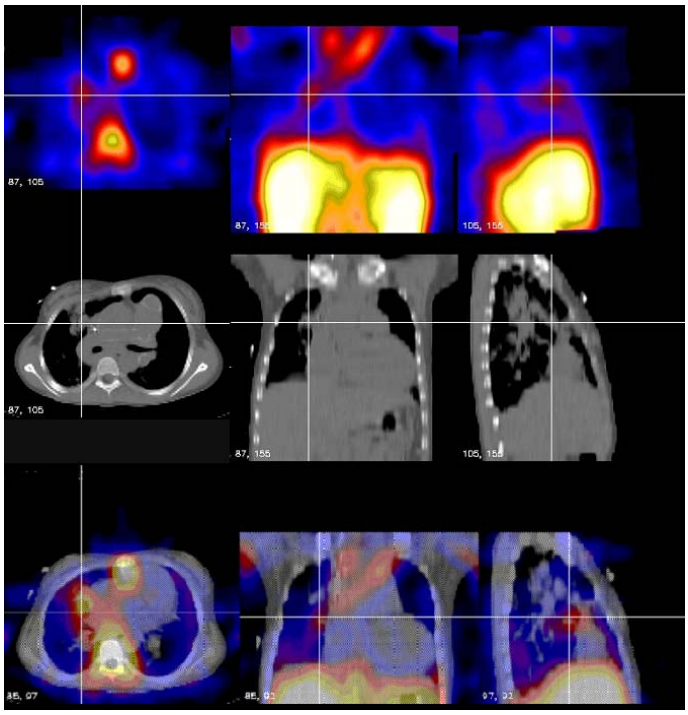


Figure 6.31: Registered gallium-67 SPET and CT showing the focal localization of uptake. The fusion of images confidently located the gallium-67 avid lesion outside of lung tissue and within a prosthetic vascular conduit. SPET (top) and CT (middle) images of three orthogonal sections (left to right: transverse, coronal and sagittal) are displayed with a linked cursor. Also shown is the fused display (bottom).

6.7 Conclusion

Registration of SPET and CT images in the abdominal and thoracic areas has been recognized as a difficult case. As described in this chapter, a digital phantom was developed and applied to simulate SPET images of different activity distributions and different radiopharmaceuticals. The simulations were used to investigate SPET-CT registration in the thoracic-abdominal region using similarity measures based on joint

intensity histogram. One aim was to determine whether the choice of a reconstruction algorithm (OSEM or FBP) affected registration performance under various conditions. Evaluation based on the SPET phantom is not affected by the quality of SPET images and avoids difficulties associated with fiducial markers. The evaluation techniques described here can be extended to incorporate local transformations in order to better simulate clinically realistic misregistration. One notable result obtained in this study is that performance of registration is relatively insensitive to image noise, irrespective of which reconstruction algorithm was used. It is also noted that the SPET-CT registration is dependent on a particular SPET image and probably application specific. No single method could be applied with equal success in all SPET applications. Results of this study also show that the registration accuracy, to some extent, depends on which algorithm (OSEM or FBP) is used for SPET reconstruction. For a given image noise level, OSEM-reconstructed SPET was found to give better registration than FBP-reconstructed SPET when AC is included, although less obvious for SPET without AC. Furthermore, attenuation compensation shows an effect on FBP SPET but little effect on OSEM SPET. These results suggest that OSEM could be a preferable choice of SPET reconstruction algorithm producing more accurate retrospective image registration when AC is used to remove artifacts due to non-uniform attenuation in the thoracic region. Registration deteriorated with an increase in projection data noise level (decrease in planar projection count). The presence of the body boundary in the SPET image and matching fields of view appeared not to be critical for improving registration, but pre-processing steps such as CT intensity windowing are helpful to improve registration accuracy. Results of experiments also demonstrated that image volume has no significant effect on registration, and registration performance deteriorated only when the volume is much smaller, such that there is too little information content to construct the intensity histogram.

Chapter 7

Development of a non-rigid registration method

In this chapter, a fully automated non-rigid image registration method that maximizes a local voxel-based similarity metric is investigated. Although the experiments described here were all carried out in 2D, the method can be readily adapted to 3D. Moreover, the method forms the basis of the 3D experiments described in Chapter 8. For these reasons, references to 3D parameters are also given, where practicable. The proposed algorithm for non-rigid registration is based on overlapping image blocks which are defined using a 3D grid pattern. The transformation vector field, representing image deformation is found by translating each block so as to maximize the local similarity measure. A symmetric version of the correlation ratio (SCR) is proposed with a median filter to ensure local coherence of the sampled displacement vectors. The resulting sparsely sampled vector field is interpolated using a Gaussian function to ensure a locally smooth transformation. A hierarchical strategy is adopted to progressively establish local registration associated with image structures at a diminishing scale. Simulation studies were carried out to evaluate the proposed algorithm and to determine the robustness of various voxel based similarity measures: mutual information, normalized mutual information, correlation ratio and SCR. In this chapter, a T_1-weighted magnetic resonance image was used to test intra-modality registration. Proton density and T_2-weighted MR images of the same subject were used to evaluate inter-modality registration. The proposed algorithm was tested on the 2D MR images distorted by known deformations.

7.1 Introduction

Section 2.3 reviews the general methodology used for image registration, which is focused mainly on rigid body alignment. This section will discuss non-rigid registration in more detail. The image registration process seeks to find an optimum spatial transformation that maps each voxel of one image into another image. Most registration methods are constrained to global rigid transformations. While useful in cases where the rigid body assumption is globally valid, they are inadequate where the assumption is valid only locally (e.g. articulated joints), where deformations are non-rigid (e.g. soft tissue organs deformed by posture changes), or where images of different subjects are

matched. Inter-subject registration may be used to enhance the diagnostic value of an image by comparing the patient's image to a control group average. Non-rigid registrations have been applied to applications such as atlas-based segmentation [353,354] and image-guided surgery [355]. In anatomical atlas matching [356,357], non-rigid registration allows automatic anatomical labeling of regions in an MR image based on their spatial correspondence to atlas features. For example, Vemuri *et al.* [353] successfully applied non-rigid registration to automatic hippocampal segmentation from MR images. Hartmann *et al.* [354] successfully quantified the brain and cerebellum volumes in normal subjects and chronic alcoholics from MR images. Shen *et al.* [358] implemented a deformable registration for constructing a statistical atlas of prostate cancer to determine an optimal biopsy strategy. Non-rigid registration in serial MR images in the brain is crucial for the neurosurgeon. Ferrant *et al.* [355] tracked intraoperative brain deformation accurately to provide critical information for neurosurgery. Rey *et al.* [359] used non-rigid registration to detect regions with volume variation in local regions by segmenting lesions in a temporal series of images. Subsol *et al.* [360] used an anatomical atlas of the skull to study craniofacial disease related to maxillary deformation. Other examples include monitoring after therapy, and differences in patient position (example is head and neck where rigid repositioning is difficult). The need to minimize local differences, common to these applications, motivated the work on non-rigid registration described in this chapter.

Several methods for non-rigid registration have been reported [47,100,361]. Most involve constrained local deformation of one image with respect to the other. Static constraints are imposed as landmarks or control points. Fitting a surface spline (such as the thin plate spline [198,203], or the elastic body spline [200] through the supplied control points yields spline coefficients. Deformation at any point in the image space can then be computed. The selection of landmarks and their matching is recognized to be a difficult and labour-intensive task if performed manually. Non-rigid registration can be formulated as a dynamic process where image warping becomes a compromise between external driving forces arising from the gradient of a similarity criterion (expressing attraction of similar features) and internal restoring forces representing the physical properties (such as rigidity) of the body. Based on this approach, Bajcsy *et al.* [199] formulated a 3D elastic image warping algorithm. In order to allow large local

deformations, a viscous fluid registration method was introduced by Christensen *et al.*[201]. Bro-Nielsen *et al.* [202] used a convolution filter to speed up the fluid registration process. Non-rigid registration can also be treated as an energy optimization problem expressed as a (Bayesian) likelihood maximization problem with a Gibbs' prior [362]. Local energy can be thought of as a prior that penalizes unlikely deformations. Gee *et al.* [363,364] formulated a non-rigid registration by modeling the cross-correlation between image intensities as an energy function. An elastic energy prior that acts to promote smoothness in image transformation penalizes deformation. Another likelihood function, based on the sum of squared differences (SSD) of voxel values between images, was proposed by Christensen *et al.* [201]. Their prior was encoded as the elastic energy of the viscous fluid and was relaxed gradually to permit large deformations. Szeliski and Lavallee [365] used SSD as the likelihood function. Deformation was constrained by zero-order and first-order continuity to regularize the transformation. Ruiz-Alzola *et al.* [366] used the intensity and gradient information to compute correlation for local matching of areas with a high degree of local structure.

Other approaches to non-rigid registration, such as the optical flow model [367], elastic model [199] and viscous fluid model [201,202,368], have been surveyed and compared in [100]. In these non-rigid registrations, the driving forces were based either on difference images or on the intensity gradients at boundaries. A major downside of these approaches is that they are restricted to intra-modality applications; otherwise, a pre-registration segmentation step is required to define surfaces of interest. An important feature of these registration algorithms is that the adopted physical model imposes significant constraints on the deformation. Another major feature of the elastic and fluid algorithms is the long computation time, even using massively parallel computers [368,369]. Moreover, most dynamic non-rigid registration methods rely on the pixel-to-pixel intensity differences between images to formulate the external force and are thus restricted to intra-modality applications. Given the limitations of the above methods, the objective of this chapter was to investigate an alternative approach to non-rigid alignment, based on free-form deformation and joint intensity histogram [186,211,370,371], which can be applied to multimodality problems for non-rigid registration.

7.2 Problem formulation

The registration problem can be posed as an estimation of the deformation of a *floating* image A with respect to a *reference* image B using transformation T parameterized by an unknown displacement vector $u(x)$ where x is the position vector in the reference image. The transformed floating image A' is defined by $A'(x) = T_u A(x) = A(x - u(x))$. A displacement vector field describes the deformation at all the image voxels. It is computationally expensive to estimate the displacement vector $u(x)$ at all positions in a 3D image. To improve the efficiency, the images are divided into overlapping image blocks centred at the nodes of a uniform 3D grid. The displacement field is sampled at the grid nodes. Using a suitable similarity measure, the optimum translational displacement is found for each floating image block with respect to the corresponding reference image. The displacement vector field is then interpolated to give the local displacement at other positions in the image. The sampling density of the displacement vector field is progressively increased. At first, a coarse grid is used to capture large-scale misregistration. Then a fine grid is used for small-scale registration. The sampling density is changed by a factor of two between adjacent levels. At each level, the displacement vector field obtained from the previous level is updated by the current estimation. The summed displacement vector field is then used to transform the original floating image. The aim of the registration is to find, for each x, a displacement vector $u(x)$ that maximizes the spatial correspondence between the transformed floating image A' and the reference image B.

7.2.1 Similarity measure

As discussed in Chapter 2 (Section 2.2), similarity measures fall into two main categories: those based on image features and those based on voxel intensity. The feature-based approach requires extracting corresponding features from the images to be registered and determining a pair-wise transformation that aligns those features. The extraction of the corresponding features from both images is generally difficult, particularly for inter-modality images. The difficulty is substantial with SPET images because of their poor resolution. The feature-based similarity measures are not considered here. The voxel-based methods generally use the full image content and operate directly on image intensity values. Image registration proceeds by optimizing a similarity measure determined from voxel values. Given that image intensities are not

linearly related in inter-modality images, the sum of squared differences and the cross correlation measures [171,363,364] are unsuitable for this study. Alternative local similarity measures include the mutual information [177,178,179,220], the normalized mutual information [180,207], the correlation ratio [181,264] and the symmetric correlation ratio [182]. These similarity measures assume that the co-occurrence of intensity values in the two images is maximized at registration. In intensity space, mutual information uses entropy to measure co-occurrence of pixel values while correlation ratio uses variance to determine dispersion of the co-occurrence of pixel values in the two images.

For a given transformation T_u, we have the image pair $A'=T_uA$ and B with intensity values "a" and "b", respectively, sampled from the same finite grey scale $\{0,1,...,M-1\}$. Registration can be described as a problem of predicting the intensity values of one image given the other image. In an optimum local match, the co-occurrence of intensity values in one image block given the observation of the intensity values in the other image block is maximized. The transformed floating mage A′ and the reference image B are locally aligned over the image block (where the local similarity measure is maximum). The local alignment is identical to the rigid body registration case except for the domain over which they are defined (i.e. the image block) and the class of transformations (i.e. translation only). As defined in Chapter 3, mutual information (Equation 3.11) and normalized mutual information (Equation 3.12) are given by

$$\mathrm{MI} = \sum_{a,b} \mathrm{P}(a,b) \log \frac{\mathrm{P}(a,b)}{\mathrm{P}(b)\,\mathrm{P}(a)}, \text{ and}$$

$$\mathrm{NMI} = \frac{\sum_{a} \mathrm{P}(a) \log \mathrm{P}(a) + \sum_{b} \mathrm{P}(b) \log \mathrm{P}(b)}{\sum_{a} \sum_{b} \mathrm{P}(a,b) \log \mathrm{P}(a,b)}.$$

The joint probability P(a,b) and the marginal probabilities: P(b) of the reference image B and P(a) of the transformed floating mage A′; can be estimated from the normalized joint and marginal intensity histograms, respectively. The correlation ratio (Equation 3.23) is also reproduced here, for the sake of convenience,

$$\mathrm{CR} = 1 - \frac{\sum_b \sum_a \mathrm{P}(a,b)\,(a - \mu_b)^2}{\sum_a \mathrm{P}(a)\,(a - \mu)^2} \; .$$

The symbol μ denotes the overall average intensity in A′. The subscripted version μ_b gives the average intensity of an induced segment in A′, which corresponds to all voxels of intensity "*b*" in B. In general, the correlation ratio is not symmetric, $\mathrm{CR}(\mathrm{B},\mathrm{A}') \neq \mathrm{CR}(\mathrm{A}',\mathrm{B})$, its behaviour depending on which image, A' or B, is used to compute the variance σ^2. A symmetric CR was proposed in Chapter 3 and defined (Equation 3.24) as $\mathrm{SCR} = \mathrm{CR}(\mathrm{A}',\mathrm{B}) + \mathrm{CR}(\mathrm{B},\mathrm{A}')$. Note that the computational efficiency of the optimization process can be improved by subsampling the image block. This and other efficiency considerations will be discussed in Subsections 7.2.2 and 7.3.1. To obtain an optimal local matching, minimization of negated similarity measures was performed. The downhill simplex method [127] was used to optimize the displacement of each image block. This method does not require the computation of derivatives or line minima. The optimization is initialised at the origin of the parameter space.

7.2.2 Interpolation of the displacement vector field

Estimates of the displacement field for image registration are generally not well defined. Outlier displacements can appear as discrete impulse-like noise. Imposing constraints to penalize unlikely deformations can circumvent this problem. For example, a limit can be placed on the maximum displacement at each grid node. However, this method may prevent the estimation of large but valid deformations. The proposed method does not impose an upper limit on the local displacement. Instead, a median filter is employed to ensure local coherence of the deformation field (Figure 7.1).

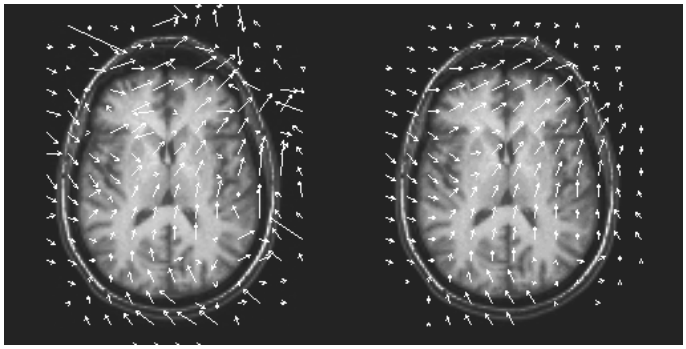


Figure 7.1: The displacement vector field before (left) and after (right) the application of the median filter.

The median filter offers superior suppression of impulse noise and generally superior preservation of detail when compared to the averaging filters, such as the boxcar or the Gaussian filter [211,372,367]. Once the displacement vectors $\tilde{v}$ have been determined at grid nodes, a 3D median filter is applied to each Cartesian component of the displacement vector field. Let the image voxels be indexed by $i \in G = \{0, 1,..., N-1\}$, and let voxels where nodes occur be indexed by $j \in Q \subseteq G$. Suppose $c(x_j)$ represents a putative displacement field such that each measured displacement vector $\tilde{v}(x_{j'})$ sampled at grid nodes j' is a Gaussian weighted sum of all $c(x_j)$. The displacement field $c(x_j)$ is used for interpolation to provide estimates of the displacement vector displacement $v(x_i)$ at intermediate locations (i.e. at an arbitrary voxel x_i) by interpolation with a Gaussian kernel (Figure 7.2),

$$v(x_i) = \sum_{j \in Q} \phi(x_{ij}) c(x_j), \tag{7.1}$$

$$\text{where} \quad \phi(x_{ij}) = \exp\left(-\frac{\|x_{ij}\|^2}{2\delta^2} \right) \tag{7.2}$$

and $\|x_{ij}\|$ denotes the Euclidean norm of $x_i - x_j$ on R^3.

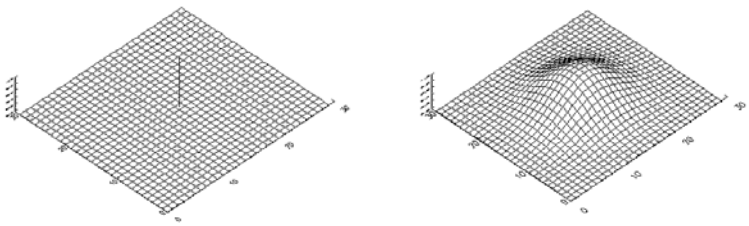


Figure 7.2: A single data sample in spatial domain (left) and after convolution with the Gaussian function (right).

To determine the coefficients $c(x_j)$ the displacement $v(x_i)$ is set equal to $\tilde{v}(x_i)$ on each of the sampled displacement data points at grid nodes $j' \in Q$. Let i in Equation 7.1 correspond to some grid node j'. A system of linear equations (Eq. 7.3) is formulated to compute c_j,

$$(\phi_{j'j})(c_j) = (\tilde{v}_{j'}), \tag{7.3}$$

where (c_j) is a column vector consisting of n coefficients corresponding to the number of nodes, $(\tilde{v}_{j'})$ is a column vector consisting of n sampled displacements, and $(\phi_{j'j})$ is a $n \times n$ matrix of Gaussian weights $\phi_{j'j}$ given by Equation 7.2. Since c is a vector in the same sense as u is, i.e. with 3 Cartesian components, Equation 7.3 is a system of three equations, one for each component. The computation of the exact solution of Equation 7.3 is CPU and memory intensive. For example, in a 3D cubic image data set of 128^3 voxels, in which every fourth voxel along each orthogonal direction forms a grid node, there are 2^{15} nodes. To improve the efficiency an approximate solution is obtained by setting $c_j = \tilde{v}_j$ (i.e. setting Gaussian δ to go to zero). In this case, the interpolation can be represented as a convolution given by Equation 7.4 of the displacement vector field sampled at the grid nodes with the Gaussian kernel.

$$\begin{aligned} v(x_i) &= \sum_j \phi(x_{ij})\tilde{v}(x_j) \\ &= \phi(x_i) * \tilde{v}(x_i), \end{aligned} \qquad (7.4)$$

where $\tilde{v}(x_i) = \begin{cases} \tilde{v}_j & , i = j \\ 0 & , \text{ otherwise} \end{cases}$

Although the Gaussian has a global support, the contributions from remote nodes are generally negligible so that a relatively small convolution kernel can be chosen thus reducing the computational cost. The effective size of the kernel is controlled by the Gaussian width δ. When δ is small, the relative influence of nearby nodes diminishes rapidly with distance from the voxel. Deformation occurs only within a small neighbourhood of each grid node without influencing the remaining image. Gaussian weighting helps to reduce computation cost because only a small subset need be considered. A practical implementation constrains the support to the nearest grid nodes giving adequate interpolation of $v(x)$ with improved computational efficiency.

7.2.3 Coarse-to-fine displacement

The transformation T_u is an image-wide transformation dependent on c, which is "piecewise" linear locally given by each block. A strategy that integrates local alignment across different coarse-to-fine level in displacement vector space offers significant advantages in terms of computational cost and registration robustness. In the

proposed implementation, the displacement vector field is first determined using large image blocks at a coarse sampling density. Then, registration proceeds by refinements through progressively smaller blocks at the finer level of displacement field (Figure 7.3).

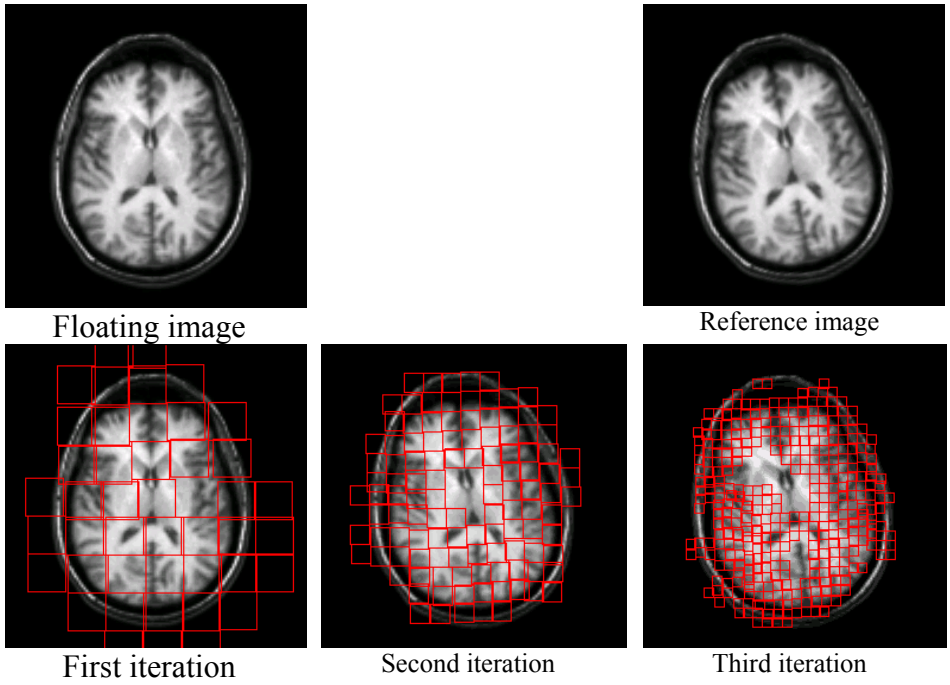


Figure 7.3: Multi-level non-rigid registration illustrated for a rotational distortion of the reference image. Superimposed on the floating image are displaced image blocks following the first, second and third iteration. For the sake of clarity, every second block is shown, and only if it is subject to a significant displacement. This explains the large gap in the middle of the 2nd and 3rd iteration results.

The low-frequency component of the misregistration error is reduced at a coarse level, whereas the high-frequency component is reduced at a fine level [365,373]. At coarse-to-fine level κ, it corresponds to iteration κ and the residual displacement field is given by Equation 7.4. The displacement $u^{(\kappa)}(x)$ transforms the floating image which is then used to estimate the residual displacement field $v^{(\kappa)}(x)$ at a more refined level of the next iteration. The previous estimate $u^{(\kappa-1)}(x)$ is added to the residual displacement field $v^{(\kappa)}(x)$ producing an updated displacement vector field $u^{(\kappa)}(x)$,

$$u^{(\kappa)}(x_i) = \begin{cases} u^{(\kappa-1)}(x_i) + \phi^{(\kappa)}(x_i) * \tilde{v}^{(\kappa)}(x_i), & \kappa > 0 \\ 0, & \text{otherwise} \end{cases} \qquad (7.5)$$

where $\phi^{(\kappa)}(x_i) = \exp\left(-\frac{\|x_i\|^2}{2(\delta_\kappa)^2}\right)$ and δ_κ denotes the Gaussian width at iteration κ.

Updating only the displacement field, and not the image itself, avoids cumulative image blurring arising from successive interpolations of the floating image. To improve the

computational efficiency, nearest grid nodes were used in the interpolation and the degree of interpolating polynomial varied with block size.

7.3 Methodology

The practical considerations of grid node spacing, image block size and sampling scheme for the multi-level implementation of the proposed non-rigid registration is introduced in Section 7.3.1. The details of the image data set for intra- and inter-modality registration studies are given in the Subsection 7.3.2. The results obtained from the experiments in this chapter will form the basis for the 3D experiments in Chapter 8.

7.3.1 Implementation

The registration proceeds progressively from coarse level to fine level with square image blocks of size 64^2, 32^2 and 16^2 pixels in 2D (64^3, 32^3 and 16^3 voxels in 3D), corresponding to grid spacing of 32, 16 and 8 pixels, respectively. The inter-block distance is chosen such that each block was completely overlapped by its 8 neighbours in 2D (see Figure 7.4) and 26 neighbours in 3D during local matching. The centres of neighbouring blocks must be located at midpoints of the sides or the corners of a block. After each translation, a tri-linear interpolation was used to obtain the intensity values in the translated image block.

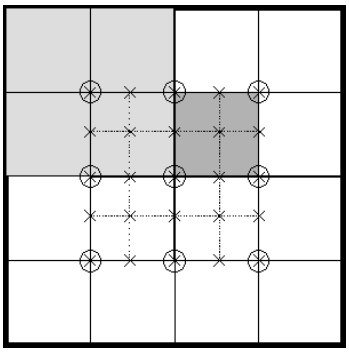


Figure 7.4: Image blocks and grid nodes are illustrated for two adjacent levels: at the coarse level, grid nodes are circled and the image block is lightly shaded. The centre block is completely overlapped by its 8 neighbours (centred at circles); at the fine level, nodes are shown with crosses and the image block is darkly shaded. Note that the coarse grid nodes persist to finer levels.

The image intensities of each image block were binned into a 2D joint intensity histogram for subsequent estimation of the similarity measure. Rebinning was performed implicitly by linear rescaling of the image intensity. To improve estimation of the joint and marginal probability distributions from the corresponding histograms,

the number of bins in the joint intensity histogram was reduced as the block became smaller [274,374]. There were 64 bins in the 64×64 block, 48 in the 32×32 block, and 32 in the 16×16 block. The number of bins was derived empirically [375]. Subsampling of the image block can be used to speed up the registration. Notation [n_1,n_2,n_3] represents down-sampling of image blocks uniformly by a factor of n_1 in the first iteration, n_2 in the second iteration, and so on. For example, a [4,2,1] sampling scheme applied to a square block of size 32^2 pixels will generate 8^2 samples in the first iteration, 16^2 in the second, and 32^2 in the third. At each level, the iterative optimization was terminated as in Subsection 4.3.5 when consecutive absolute changes in the similarity measure fell below 0.0001. Failing that, the algorithm stopped when a certain number of iterations were reached (50, 40, and 30 in the first, second and third level, respectively). Fewer optimization iterations were allowed at finer levels where the computation was more expensive. All studies were carried out on a dedicated 333 MHz Pentium PC. The setting mentioned in this paragraph together with SCR is referred to as the default implementation in the rest of this chapter unless otherwise specified.

7.3.2 Image data

To assess the subsampling with rebinning in the context of the multi-level implementation of the local registration algorithm, five pairs of T_1 and T_2 MR images were used, one of which is shown in Figure 7.5 (top). The images in each pair were acquired at the same time, showed no visible manifestations of misalignment and were therefore assumed to be registered. A 2D transverse section was extracted from each pair for this experiment. All 2D transverse images consisted of 256×256 pixels, each of size 0.9×0.9 mm^2, with 8 bit encoded intensity values. For this purpose, an image A is distorted elastically using a thin-plate spline [198] with 5 control points, as illustrated in Figure 7.5 (bottom). The four corner points are fixed, while a point in A was randomly selected and distorted with a displacement of random magnitude (within predetermined limits) and direction.

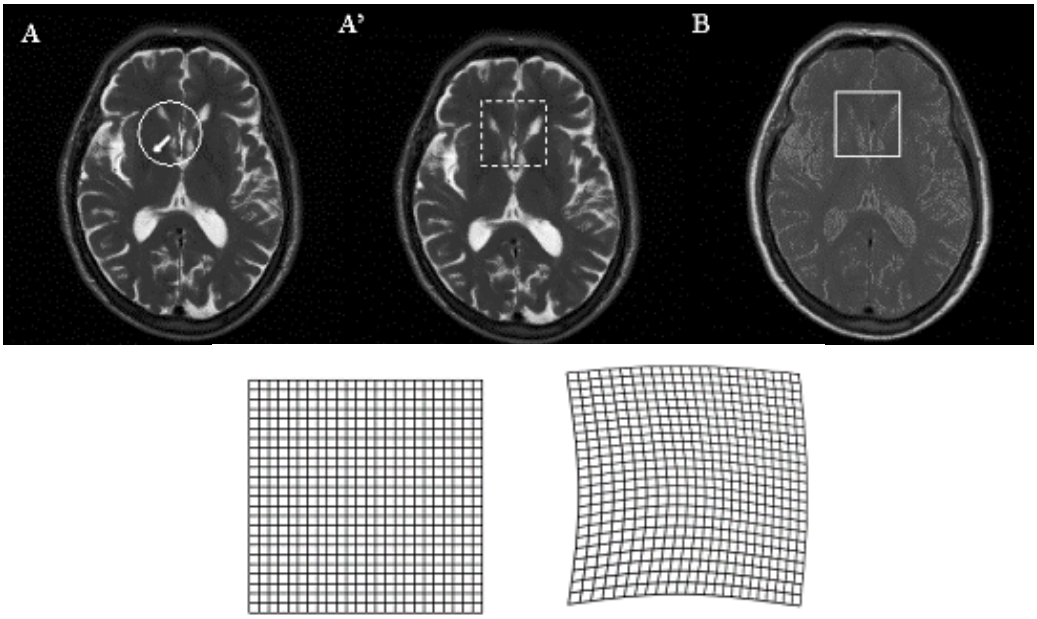


Figure 7.5: An example of data used for the evaluation of the multi-level registration strategy. An MR T_2 image A (left) is elastically transformed with 5 control points: fixed vertices and an interior point subject to random displacement within a circle of 15-pixel radius (shown). The transformed image is denoted A′ (middle). The transformation is also illustrated with the use of the image grid (bottom). An MR T_1 image B (right) was acquired at the same time as A. An example of a block used to recover the distortion is outlined by a square. Image B is registered to image A′.

To evaluate the proposed algorithm for local registration, two 2D MR image data sets (simulating intra- and inter-modality problems) were used. The intra-modality image set was based on a T_1-weighted MR image (0.937 mm/pixel) of size 256^2 pixels (Figure 7.6a). Five analytically known transformations were employed to produce deformations in this T_1 MR image and the deformed image was taken as the reference image. The synthetic deformations were bilinear, rotational, sinusoidal, thin plate spline [198] and Gaussian [376]. The original T_1 MR image was used as the floating image to be registered to the deformed image.

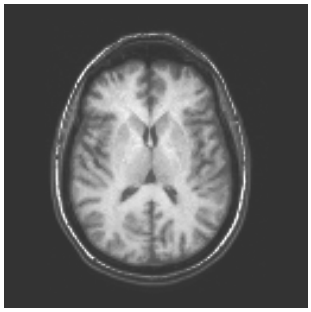


Figure 7.6(a): A T_1-weighted MR image (left) used for synthetic intra-modality registration.

The inter-modality image set, was based on a pair of T_2-weighted and proton density 2D MR images (0.898 mm/pixel) containing 256^2 pixels of the same subject (Figure 7.6b).

The two images are assumed to have been registered at acquisition. The T_2 image was deformed by the same synthetic deformations as those applied in the intra-modality case and taken as the reference image. The proton density image was used as the floating image to be registered to the reference image.

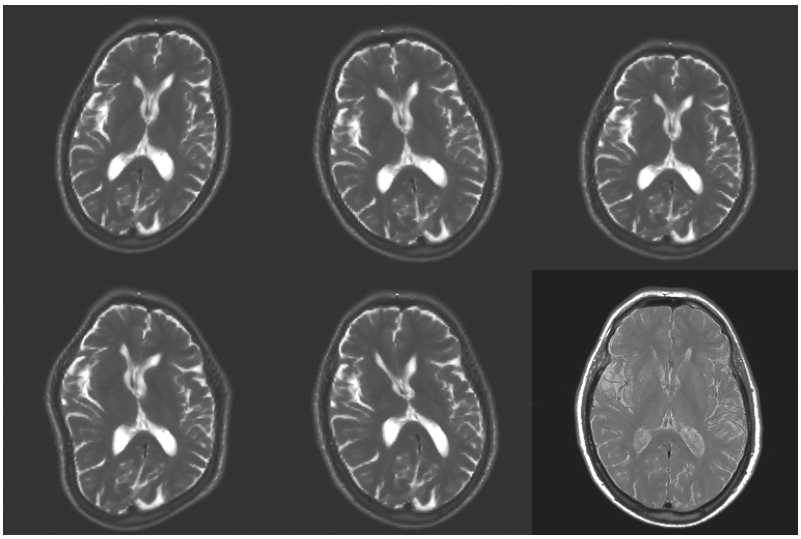


Figure 7.6(b): For inter-modality registration, synthetically deformed T_2-weighted MR images acted as the reference. Five types of deformation were used (counterclockwise from top right): thin plate spline, rotational, bilinear, rotational, sinusoidal, and Gaussian. A proton density MR image (bottom rightmost) was taken as the floating image.

7.3.3 Registration error analysis

To assess the performance of the proposed non-rigid registration algorithm, registration accuracy for synthetically deformed images was measured with respect to the known displacement field. For each voxel, error $\varepsilon_i = \|u_i - \tilde{u}_i\|$ measured the discrepancy between the applied displacement u_i and the displacement $\tilde{u}_i$ recovered by the registration process. Squared and summed over all voxels in the image, this gives the root mean square displacement error,

$$\varepsilon_{rms} = \frac{1}{\sqrt{N}} \sqrt{\sum_{i=0}^{N-1} {\varepsilon_i}^2} ,$$

where N is the image size. In 2D registrations, the image was masked to exclude displacement falling outside the reference image B or the floating image A′. For consistency ε_{rms} is referred by "DE" in the rest of this chapter. DE values are expressed in pixels (of size 0.937 mm). As a measure of the performance of registration, DE averaged over all synthetic images was also recorded. Statistical analysis (two-tailed paired t-test) was applied to differentiate the average DE obtained by two registration

algorithms. Studies by independent researchers suggested that a 2 mm misalignment is at the limit of a trained observer's ability to detect a misalignment in intra-subject inter-modality brain registration for most clinical data [131,208,224,248]. To assess the robustness of the proposed algorithm for local registration, thus registration was considered successful if the registration error did not exceed 2 pixels (1.8 mm). Robustness of the algorithm was quantified as the proportion of successful matches in each set of experiments.

7.4 Evaluation of non-rigid registration in 2D

The non-rigid registration algorithm proposed in this chapter is formulated by using multi-level strategy with convolution to interpolate displacement field. A Gaussian kernel was convolved with the sampled displacement field to estimate a continuous and smooth transformation T_u over the data volume. This is a considerably more difficult problem than the rigid body registration. Factors such as the width of the Gaussian kernel, treatment of displacement outliers, image block size, multi-level strategy, choice of the similarity measure, and the choice between the full (Eq. 7.1) and the fast implementation (Eq. 7.4), can all influence the performance of the algorithm. Investigations were performed to assess the influence of these parameters on the non-rigid registration and to optimize the parameter values where possible. Before turning the attention to real 3D clinical image data (Chapter 8), 2D image data with synthetic distortions were used in an attempt to understand the proposed algorithm under various settings of the implementation. These results will be used to refine the 3D implementation.

7.4.1 Effects of downsampling – an experimental study

The purpose of this experiment was to investigate the effect of downsampling the image block and the number of intensity bins on the robustness of the inter-modality registration. The general motivation for image block downsampling was to accelerate registration without decreasing its performance. The robustness of registration depends on the propensity of the optimization algorithm to become trapped at local maxima in the similarity measure. For a given similarity measure, robustness depends on the type of image, downsampling factors and number of intensity bins. It is not clear how these settings would affect the performance of registration. The experiments described here will provide an assessment of multi-level image registration, where downsampling

strategies are used, and help optimize the settings for both the global and local image registration.

The selected interior point in image A (see Figure 7.5) was displaced from its original position randomly in the range 0-15 pixels with uniform distribution. The TPS transformed image was taken as the reference image and denoted by A′. B was referred to as the floating image. The transformation of A′ was non-rigid which was determined by the TPS interpolation. The floating image B was then registered to the distorted image A′ using image blocks to recover the displacement. For each of the 5 image pairs, the TPS distortion and the image block registration was repeated for 10 random displacements of each of the 20 randomly selected control points, giving a total of 1000 registrations for each experiment. The experiment was carried out for each similarity measure (MI, NMI, CR and SCR) and each of three block sizes (64×64, 32×32 and 16×16), with 128, 64, 48, 32 and 16 intensity bins as well as downsampling by a factor of 2, 4 and 8 along each axis. For each experiment, the registration error and its distribution were recorded. Since subsampling is used synonymous with downsampling, subsampling refers to downsampling in the rest of this chapter for the sake of consistency in the thesis.

7.4.2 Effect of the Gaussian kernel width

To examine the influence of the width δ of the Gaussian kernel on registration, the intra-modality image set was used. Default implementation (Subsection 7.3.1) was used for the local match. Values of δ ranging from 0.3 to 1.0 of the grid spacing were used. The displacement error was averaged over the five synthetic deformations.

7.4.3 Effect of other settings

To evaluate specific aspects of the proposed algorithm, five distinct implementations were distinguished (Table 7.1). The intra-modality image set was used. DE and average DE values over the five distortions were computed for each setting at every iteration level to study the effect of setting on registration. In the full implementation, coefficients c_j used in interpolating the displacement field were determined from exact solutions of (7.3).

Implementation	(i) TPS	(ii) Unfiltered	(iii) 32^2 block	(iv) Full	(v) Fast
Interpolator	Thin-plate	Gaussian	Gaussian	Gaussian	Gaussian
Support	all nodes	3×3 nodes	3×3 nodes	all nodes	3×3 nodes
Median filter	Yes	no	yes	yes	yes
Multi-level	Yes	yes	no	yes	yes

Table 7.1: Details of five implementations. In the multi-level algorithm, blocks of size 64^2, 32^2 and 16^2 pixels were used with 64, 48 and 32 bins, respectively. The interpolator support was confined to a 3×3 neighbourhood of grid nodes except for the "full" and the TPS implementations. All implementations used the SCR similarity measure and image blocks were fully sampled except for (v) which used the subsampling default implementation described in Subsection 7.3.1. 3×3 nodes refer to the number of neighboring grid nodes used for interpolation.

7.4.4 Effect of filter type on registration

Three low-pass displacement field filters were examined: two averaging filters (boxcar and Gaussian) and a median filter. The registration of the T_1-weighted MR image with the synthetically deformed images was carried out with the set of parameters specified in Section 7.3.1. In the Gaussian filter, the width δ was set equal to the grid cell size. A preliminary study found that a 3×3 kernel (3×3×3 in 3D) was of optimum size for the filter, and this was used as the default. DE and average DE over the five distortions were computed for each setting at every iteration level. For each filter, implementations in both Cartesian and polar coordinates were carried out. In polar coordinates, the filter was applied separately on the vector magnitude (radial component) and orientation (phase angle) while, in Cartesian coordinates, the filter operated separately on each Cartesian component.

7.4.5 Comparison of similarity measures

To assess the performance of SCR relative to other similarity measures (MI, NMI, CR1 and CR2) in the intra-modality non-rigid registration, the T_1-weighted MR was registered to the five synthetically distorted T_1-weighted MR. The default implementation was used (Section 7.3.1). Four sampling schemes were applied: [1,1,1], [2,2,1], [4,2,1] and [8,4,1]. SCR was also compared to the other similarity measures in the inter-modality non-rigid registration. The proton density MR was registered to the five synthetically distorted T_2-weighted MR using the same implementation as for the intra-modality registration.

7.5 Results of evaluation of non-rigid registration in 2D

Results of various subsampling schemes and evaluation of the proposed algorithm for local registration were given in this Section.

7.5.1 Effects of downsampling

Figure 7.7 show the robustness measured as the percentage of successful matches ($\varepsilon_{rms} \leq$ 1.8 mm) for a number of subsampling regimes. Plot (a) shows variation of robustness with the subsampling factor (128 bins in all cases) and plot (b) shows variation with the block size for a given subsampling factor. Plots (c-h) show variation with the number of intensity bins in the joint histogram for specified similarity measures. The following conventions are used: n_x and n_y refers to subsampling factor in x- and y- direction, M refers to the number of intensity bins. $N_x \times N_y$ denotes the size of images. In plot (h), CR1 and CR2 (Eq. 3.23) refer to two cases of normalizing conditional variance using two different images (e.g. T_1 and T_2 MR). The general observations of the results are:

1. Registration in which the image blocks are fully sampled performs, in most cases, better than that with sparse sampling regardless of image block size, the number of intensity bins and similarity measure. An example is illustrated in Figure 7.7(a) using SCR. In this case 69%, 62% and 36% robustness was found using block size of 32×32 for [1,1], [2,2] and [4,4] sample scheme was used, respectively. This is expected as sparse sampling reduces the statistical power of the joint histogram. However, subsampling affected the performance of SCR to a lesser extent compared to other similarity measures. For example, subsampling a 64×64 image block by a factor of 2 (74%) and 4 (73%) when compared to full sampling (69%) did not have a substantial effect on SCR robustness, irrespective of the number of intensity bins (Figure 7.7(f)-(h)).

2. Robustness was found to improve with larger image blocks suggesting that more information content provided by the larger block improves statistical power for matching. However, the improvement was not uniform and depended on the similarity measure used for the alignment. The registration deteriorated with decreasing block size irrespective of the similarity measure, but SCR was least affected. As an illustration, in Figure 7.7(b), SCR is compared to MI for three block sizes. The results demonstrate that the SCR is more robust than MI for image alignment.

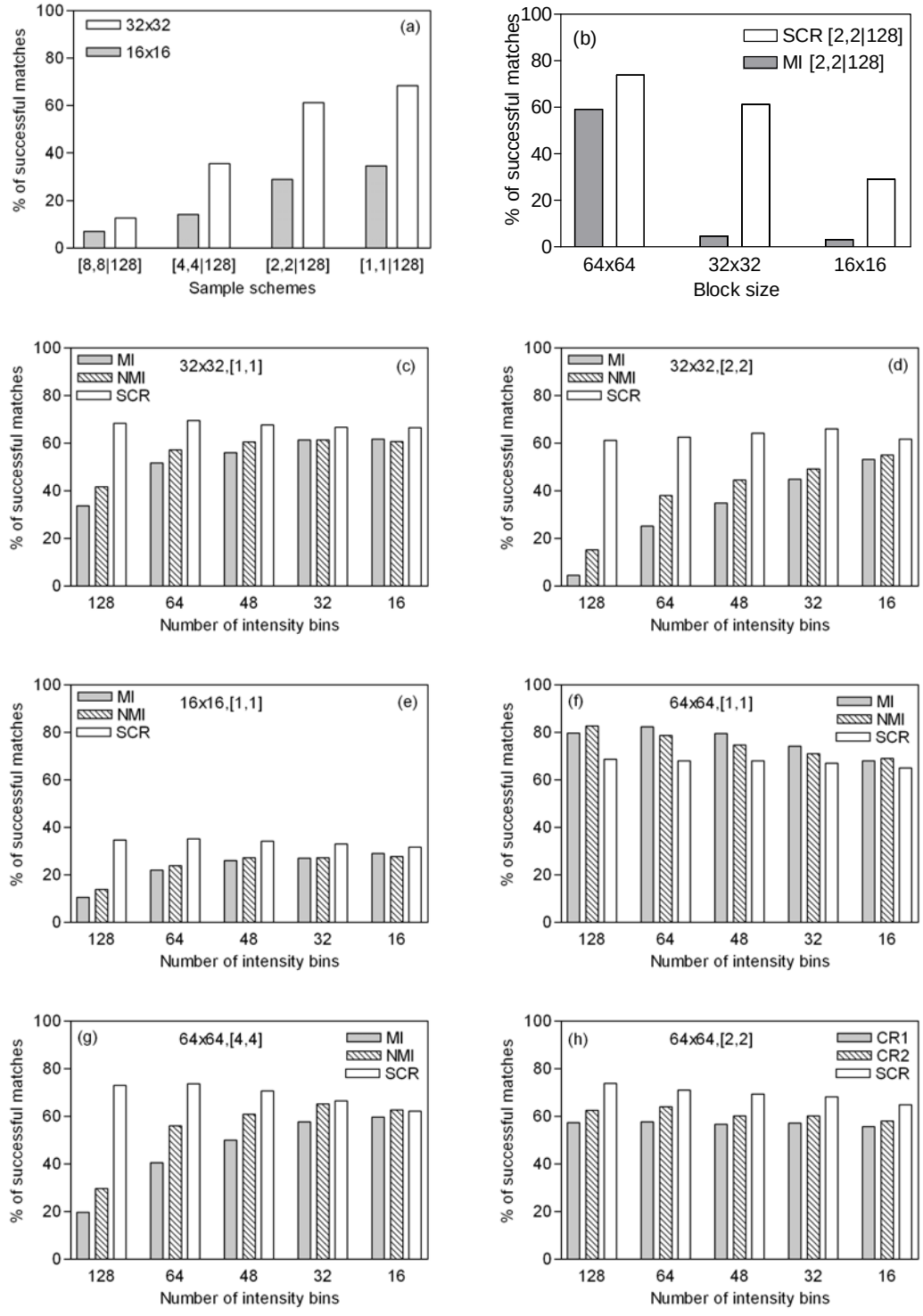


Figure 7.7: Percentage of successful matches (out of 1000 registrations) for indicated similarity measures is plotted against (a) downsampling factor; (b) block size $N_x \times N_y$. The results obtained when using full sampling and 128 intensity bins are included as a reference in (a). (c)-(e): Percentage of successful match of MI, NMI and SCR is plotted versus the number of intensity bins *M*. (f)-(g): Percentage of successful match of MI, NMI and SCR is plotted versus the number of intensity bins *M*. (h): Percentage of successful match of CR1, CR2 and SCR is plotted versus the number of intensity bins *M*.

3. Registration performance varied with the number of intensity bins but the variation depended on the choice of the similarity measure. Figure 7.7(c)-(e) summarize

the registration results for the three measures: MI, NMI and SCR. For MI and NMI, robustness improved with fewer bins. At [1,1] full sampling and [2,2] subsampling with 32×32 block size, MI and NMI show a dependence of performance on number of intensity bins. The performance for SCR was almost unaffected by the number of bins regardless of block size and subsampling. The difference between the success rate for SCR and that for MI and NMI was more pronounced with finer quantization (128 and 64 bins). When 64×64 was used without subsampling, MI and NMI showed slightly better performance than SCR (Figure 7.7(f)). However, when the 64×64 block was subsampled by a factor of 4 to produce 16×16 samples, SCR performed much better than MI or NMI (Figure 7.4(g)).

4. Figure 7.7(h) compares the robustness of SCR to that of its two components (Equation. 3.23) here denoted CR1 and CR2 for the case of the 64×64 block subsampled by half in each dimension. In CR1, the T_2 image was registered to the T_2 image, and vice versa in CR2. The success rate of SCR shows a modest but statistically significant advantage over its two components and that was also the case for other block sizes and subsampling factors. Based on the data in (h), the average (±SD) success rate 70±3% of SCR was better than both CR1 (57±1%) and CR2 (61±2%) at the 5% confidence level using two-tailed paired t-test.

In the second set of experiments, registration algorithms were tested against a set of distortions. The warped images are shown in Figure 7.6. In addition, the distortion is expressed as a histogram of pixel displacement (Figure 7.8). Together, these distortions present histogram of displacement distributions. Based on the synthetic distortions, the effect of various parameter values and different of implementation settings on non-rigid registration was investigated. The dependence of the registration accuracy on the Gaussian kernel width, the choice of the interpolator, and the choice of the similarity measure were evaluated using T_1-T_1 and T_1-proton density image pairs. The effects of subsampling the image blocks were also studied.

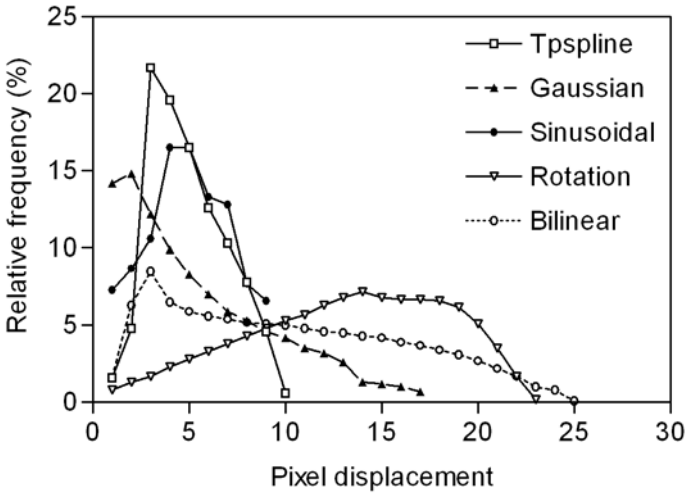


Figure 7.8: Histogram of pixel displacement for the indicated types of synthetic deformation. The deformations applied to the reference image are measured. Not included are the background pixels: 37% (TPS), 43% (Gaussian), 40% (sinusoidal), 37% (rotational) and 39% (bilinear).

7.5.2 Effect of Gaussian kernel width

The displacement error, averaged over the five synthetic distortions, is plotted as a function of Gaussian width δ (Figure 7.9). The error bars correspond to the standard deviation of the distribution of DE values across the five synthetic distortions. The result shows a decrease of ADE and gives a minimum when Gaussian width was about 0.45 of the grid spacing, but ADE deteriorates with further increase in kernel width. This value was used in all subsequent simulation studies.

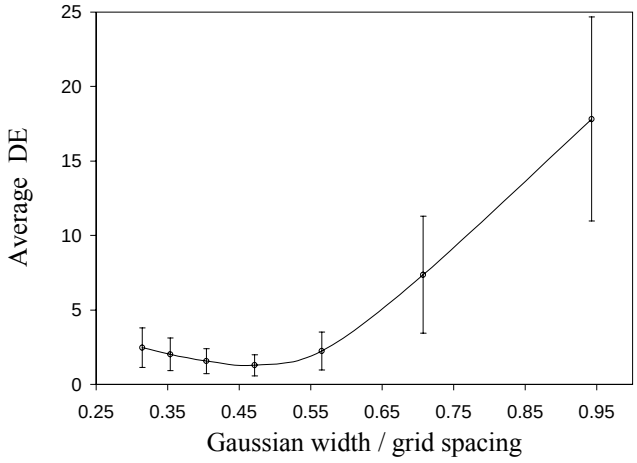
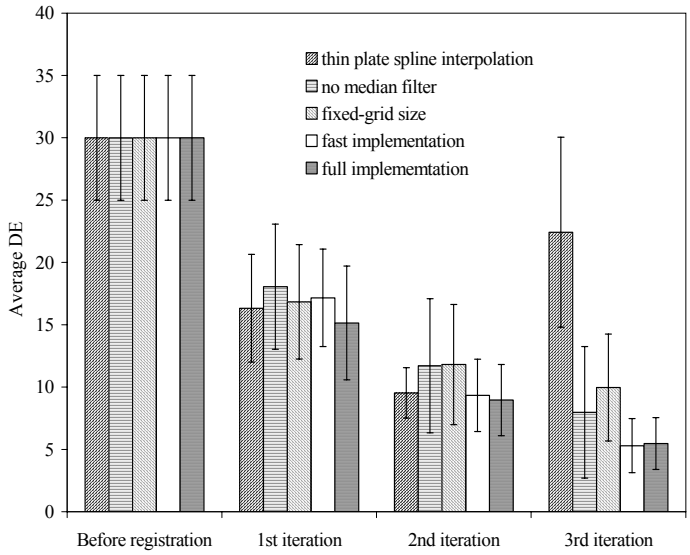


Figure 7.9: DE averaged over five synthetic deformations as a function of Gaussian width. The error bar indicates the standard deviation of the DE. Fast implementation (with 3×3 interpolator support) and full sampling of the image blocks were used for the registration. A median filter was applied to the sampled data before interpolating the displacement field.

7.5.3 Effect of settings on registration

Figure 7.10 manifests a broad trend for the average displacement error (ADE) to decrease with increasing iterations for all implementations, as expected. One exception was the implementation incorporating the thin-plate spline interpolation, which showed gross misregistration for all five distortions after the second iteration, and ADE increased after the third iteration. Using the two-tailed paired t-test, ADE was found to be significantly smaller (P<0.05) in the fast implementation than in the thin-plate spline interpolation method. The fast method also showed the smallest final ADE. Since ADE is a global measure of misregistration error, averaged over multiple deformations, it may fail to detect all local misalignments. The average DE (Figure 7.10) was found to exceed 5 pixels after three iterations.



`**Figure 7.10**: Variation of average DE with iteration number for specified implementations. Error bars indicate the standard deviation of DE, expressed in pixels (of size 0.898 mm).

This is probably due to unrealistically large synthetic distortions. Furthermore, the large error bars in Figure 7.10 reflect individual cases of significant misregistration. Examples are shown for the case of the bilinear distortion in Figure 7.11. Severe local misregistrations are observed in the absence of the median filter and where the block size is kept fixed. The results signify that the proposed algorithm (both the full implementation and the fast implementation) give superior registration. In terms of accuracy, the fast implementation is comparable or even slightly better than the full implementation. In addition, the fast strategy is more computationally efficient, requiring 0.63 minutes compared to about 3.62 minutes (for the full implementation) to

complete the three iterations (Table 7.2). The greatest gain occurred in the final iteration where the finest sampling was required. The 6-fold reduction in the overall computational cost was achieved without trading off registration accuracy.

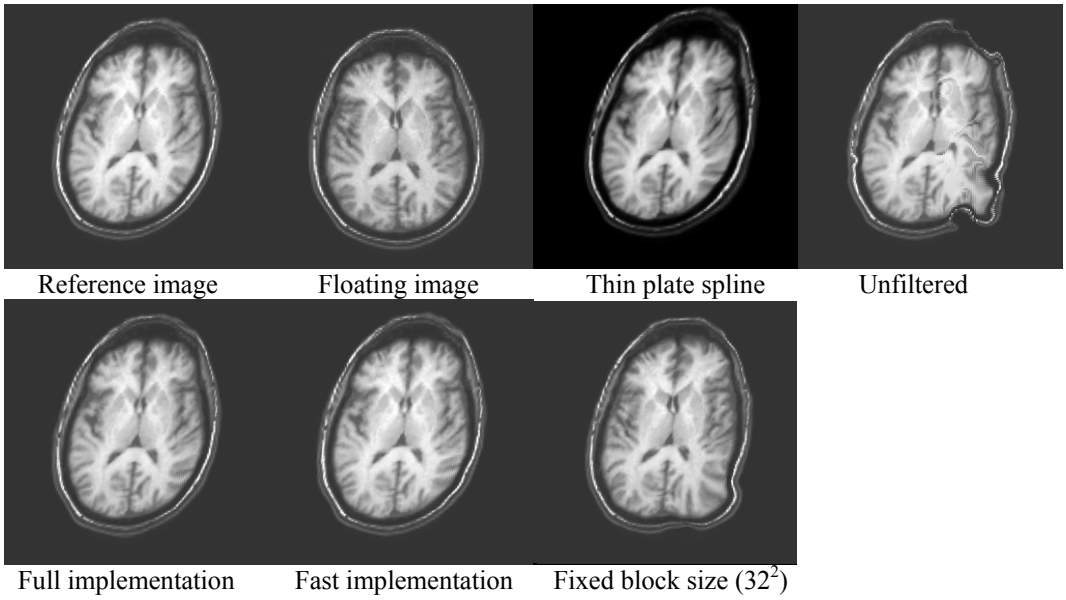


Figure 7.11: Results of registration obtained with various implementations as indicated. The reference image was distorted using the bilinear transformation. Each image is scaled independently.

	1st iteration	2nd iteration	3rd iteration	Total
Full implementation (min)	0.28	0.85	2.49	3.62
Fast implementation (min)	0.21	0.21	0.21	0.63

Table 7.2: The computational time (in minutes) for the full and the fast implementations of the proposed algorithm, averaged over five synthetic studies. See Table 7.1 for details of the implementations.

7.5.4 Effect of the filter type on registration

Performance of various filters and different implementation strategies was compared at each iteration. When a filter was applied in polar coordinates (for radial magnitude and phase angle), the median filter was superior to boxcar and Gaussian filters. Using boxcar filter and Gaussian filters in polar coordinates, misregistration deteriorated after the first iteration. The filter applied in Cartesian coordinates was more effective than applied using the polar coordinates, (Figure 7.12). The median filter operated separately on Cartesian components, in particular, showed a significantly better performance (two-tailed paired t-test, $P<0.05$) than filters operated on polar coordinates that works separately on radial magnitude and phase angle. In the Cartesian case, all filters showed a decrease of average DE with iteration. The median filter gave a smaller average DE than boxcar filter and Gaussian filter after three iterations.

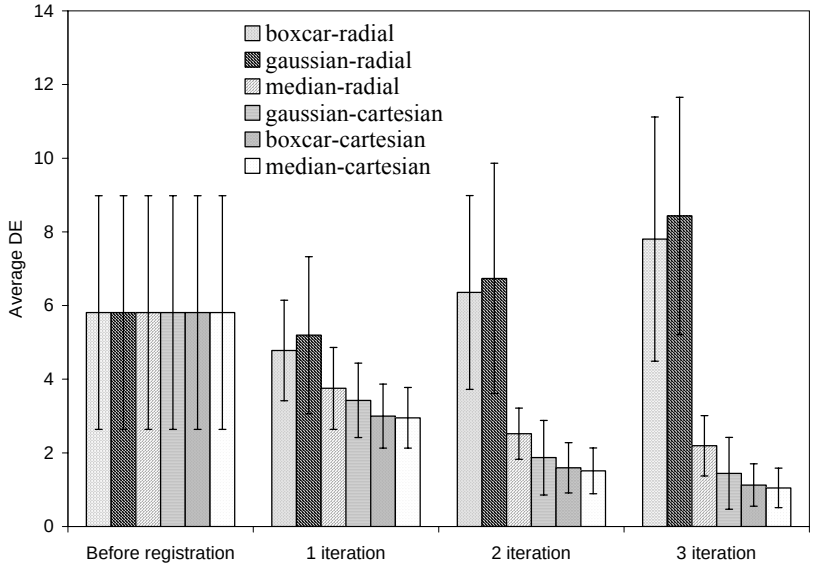


Figure 7.12: DE averaged over five synthetic deformations for different filters. The error bar indicates standard deviation of the DE. Fast implementation but with full sampling of the image blocks were used for the registration. Various filters were applied to the sampled data prior to interpolation of the displacement field. In polar coordinates, the filter was applied separately to the vector magnitude and orientation while, in Cartesian coordinates, the filter operated separately on each Cartesian component.

7.5.5 Comparison of similarity measures in intra-modality registration

In order to compare the performance of SCR to other local similarity measures, DE values were obtained for various Subsampling schemes (Figure 7.13).

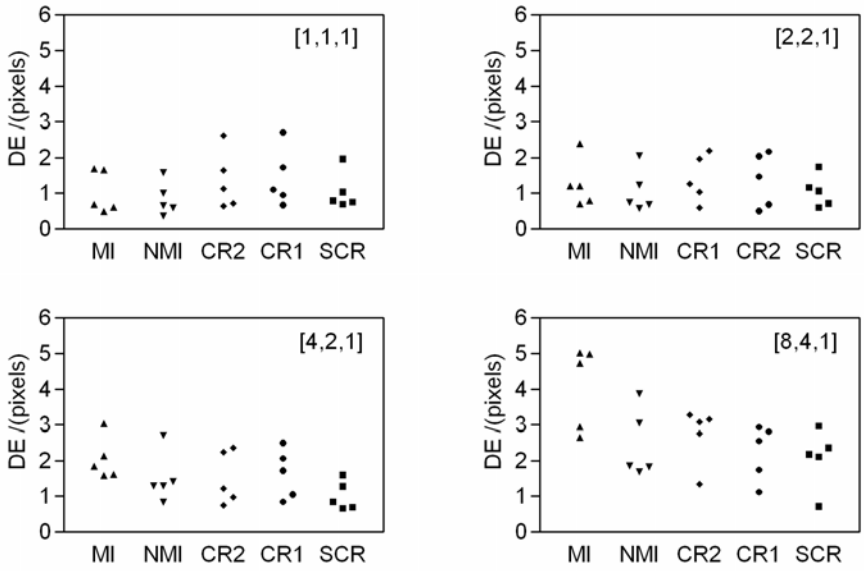


Figure 7.13: Dependence of intra-modality DE on the choice of the subsampling scheme shown for the specified similarity measures: mutual information MI, normalized mutual information NMI, correlation ratio CR1, CR2 and symmetric correlation ratio SCR.

The performance of the registration was statistically indifferent to the choice of the Subsampling scheme up to [2,2,1]. In [4,2,1] SCR shows a significantly better

registration (two-tailed paired t-test, $P<0.05$) than MI. This further confirmed the results in Subsection 7.5.1 (Figure 7.7) that SCR was more robust than MI for image block alignment with decreasing block size, irrespective of the subsampling factors and the number of intensity bins. The subsampling scheme [4,2,1] offers computational gain of about 19 *per cent* over [1,1,1] (Table 7.3). In order to compare SCR and MI in more detail, DE values were tabulated for [1,1,1], [4,2,1] and [8,4,1] sampling schemes over the various distortions (Table 7.4).

	[1,1,1]	[2,2,1]	[4,2,1]	[8,4,1]
Average time (min)	2.41	2.04	1.96	1.85
Relative time	100%	84%	81%	76%

Table 7.3: Mean processing time for registration with SCR and the specified sampling schemes, as measured on 333 MHz Pentium PC.

Distortion	DE before registration	SCR [1,1,1]	SCR [4,2,1]	SCR [8,4,1]	MI [1,1,1]	MI [4,2,1]	MI [8,4,1]
bilinear	11.0	2.0	1.6	2.4	1.7	1.8	2.7
rotational	6.1	0.8	0.8	3.0	1.7	3.0	4.7
thin plate	5.6	0.8	1.3	2.1	0.6	1.6	5.0
sinusoidal	3.0	1.0	0.7	2.2	0.7	1.6	3.0
Gaussian	3.4	0.7	0.7	0.7	0.5	2.1	5.0
Average±SD	5.8±3.2	1.0±0.5	1.0±0.41	2.1±0.8	1.0±0.6	2.0±0.6	4.1±1.2

Table 7.4: Comparison of DE values for intra-modality registration with SCR and MI under indicated sampling schemes. DE values are given in pixels (of size 0.937 mm).

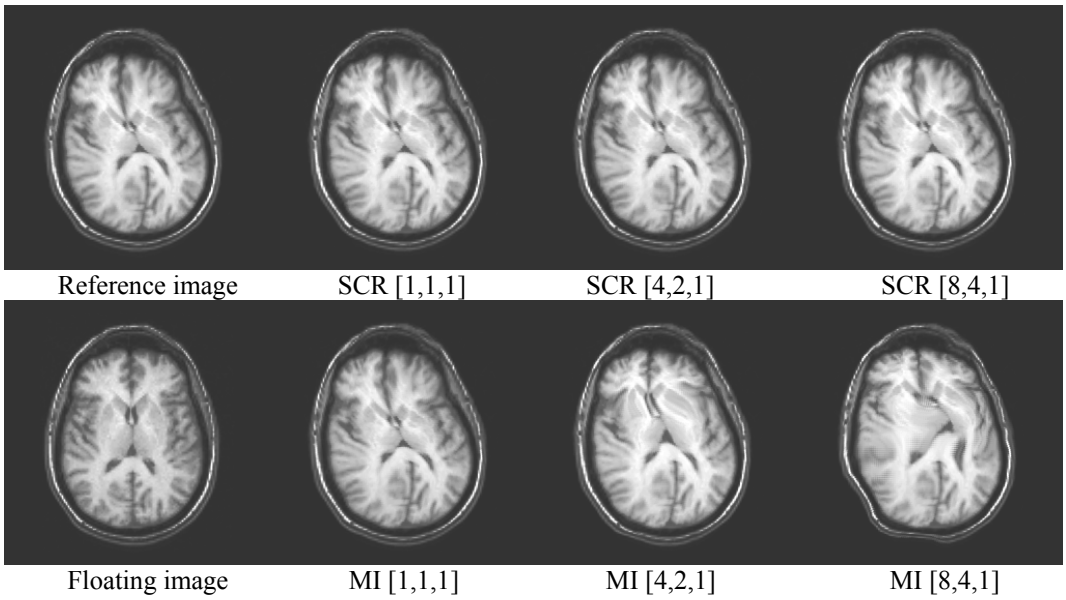


Figure 7.14: Simulation of intra-modality registration. The reference image was deformed by Gaussian distortion. Shown are the effects of the sampling scheme for two similarity measures, SCR and MI.

In addition, by way of an example, images registered to a reference image distorted by an analytical function (e.g. Gaussian) are shown in Figure 7.14. The examples illustrate the difference in performance between SCR and MI with the three Subsampling

schemes. With the sparsest [8,4,1] sampling, MI gives rise to a severe local misregistration whereas misregistration with SCR is less obvious. The discrepancy between MI and SCR persists with denser sampling [4,2,1] (see also Table 7.4). SCR clearly demonstrated superior performance to MI, in this particular case of the Gaussian distortion. The most aggressive Subsampling [8,4,1] led to a degradation of performance for all the similarity measures. However, the degradation was not uniform. MI and CR2 suffered most while CR1, NMI and SCR were least affected. The results illustrate that registration accuracy depends on the similarity measure when sparse sampling is implemented. Based on the average DE, registration deteriorated less with SCR when compared to the other methods under sparse sampling.

7.5.6 Comparison of similarity measures in inter-modality registration

Using the inter-modality data set, the performance of SCR was compared to that of other local similarity measures by measuring the pixel displacement error for various Subsampling schemes (Figure 7.15).

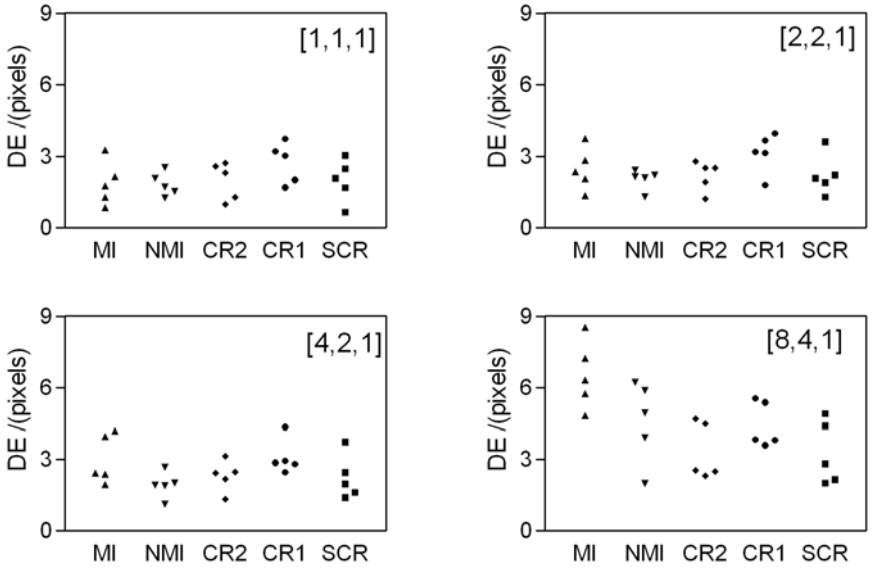


Figure 7.15: Comparison of similarity measures in inter-modality simulation experiments. DE is shown for several subsampling schemes.

The results reveal that registration performance depends on the choice of the similarity measure. CR1 had a larger DE than the other similarity measures under [1,1,1] and [2,2,1]. Degradation in registration was observed, as expected, in [4,2,1] and [8,4,1] sampling schemes but was not uniform for all the similarity measures. SCR yielded a significantly better (two-tailed paired t-test, $P<0.05$) registration than MI and CR1 in

[4,2,1]. Under the most aggressive sampling of [8,4,1], MI, NMI and CR1 deteriorated more than SCR, and their differences between ADE were statistically significant at 5% level (two-tailed paired t-test). Examples of registration results are shown in Figure 7.16. Here, the registration process attempts to correct for the rotational distortion. Under [4,2,1] and [8,4,1] sampling, local misregistration is less obvious for SCR than for MI, indicating that SCR is more robust under sparse sampling. A comparison of DE values between SCR and MI for each of the synthetic distortions is presented in Table 7.5. The registration accuracy is, on average, reduced with the sparser sampling scheme but the increase in DE is less marked with SCR (12%) than with MI (61%). In fact, sparser sampling improves accuracy in 3 out of 5 cases with SCR, but in none of the cases with MI.

	DE before	SCR		MI	
Distortion	Registration	[1,1,1]	[4,2,1]	[1,1,1]	[4,2,1]
bilinear	11.0	2.1	2.0	1.8	2.0
rotational	6.1	3.0	2.4	3.3	4.2
thin plate	5.6	0.7	1.6	0.9	2.4
sinusoidal	3.0	2.5	3.7	2.1	4.0
Gaussian	3.4	1.7	1.4	1.3	2.4
Average±SD	5.8±3.2	2.0±0.9	2.2±0.9	1.9±0.9	3.0±1.0

Table 7.5: Comparison of SCR and MI for inter-modality registration [1,1,1] and [4,2,1] sampling schemes. DE, in pixels (of size 0.898 mm) was computed between the registered T_2-weighted MR image and the corresponding transformed T_2-weighted MR image for a range of synthetic distortions.

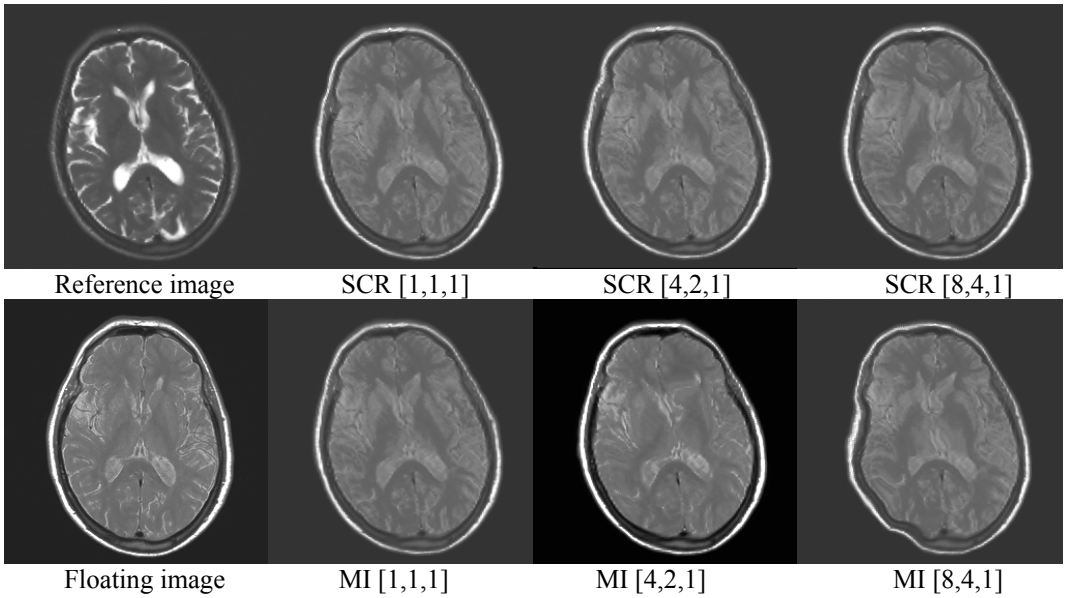


Figure 7.16: Simulation of inter-modality registration. The reference image was globally rotated. Shown are the effects of changing the sampling scheme. Two similarity measures, SCR and MI, are compared. Images are scaled independently.

7.6 Discussion

The proposed algorithm (both the full implementation and the fast implementation) gave better registration than the thin-plate spline implementation when a fine sampling grid was used. The thin-plate spline interpolation gave an optimal registration at a large grid spacing of 32 pixels, and registration deteriorated with finer grids. A similar finding by Kostelec *et al.* [377] reported the optimal spacing of 16 pixels. When compared to global registration, the use of image blocks to define local matching is likely to increase noise in the similarity measure and possibly introduce local minima. It is therefore to be expected that the misregistration will increase. One could take the view that whenever DE is not zero, there is a misregistration, giving misregistration rate of 100%. As such, it will be sensible to define a success rate (robustness), which is based on a predetermined misregistration, to compare performance of registration over various parameters or settings. In this case, 1.8 mm misregistration (Subsection 7.3.3) was used as the threshold to compute robustness of registration.

The proposed non-rigid registration uses a multi-level approach where image block size is progressively reduced to capture coarse-to-fine residual displacement vector field. The image block is downsampled successively with the same sampling interval. So while displacement field interval is the same in image pixels, it increases in mm, and the residual displacement vector fields are combined with downsampling the blocks except that no blurred version of image block was created for each level before registration. In this case, image blocks are matched directly without blurring. The general strategy of the proposed registration, strictly speaking, is not a multi-resolution match because the image is not scaled, and it's probably not resolution of displacement field that changes as registration progresses between levels. Only block size varies, the effect of which is subtler and harder to describe. This can be described as follows: Suppose v_0 gives a true displacement for each point in the image space. Replacing a point with a voxel-sized block substitutes an average intensity value in the voxel for the delta-sample. Because of its finite size, the voxel can be rotated and warped, not just translated. Therefore, the measured displacement $\tilde{v}$ is only an approximation of the true transformation at that point. Suppose now the block contains multiple voxels. Each voxel has its v so the result will be some compromise among competing displacements, which is not necessarily an average; it will depend on the similarity measure. The bigger the block the poorer the approximation to the true displacement v_0 at its centre (or any other point in the block).

But, if the true transformation is a translation uniform over the block, the similarity measure will deliver a "strong" peak at registration, with good accuracy. However, if other transformations (e.g. rotation about block centre) dominate, approximation will be so poor that the similarity measure may not register a significant peak anywhere. So long as the underlying field is uniform translation, increasing the block size increases the "power" and accuracy of registration. If the assumption is not valid, reducing the block size will improve accuracy. Pluim *et al.* [275] compared direct (without blurring) and multi-resolution matching for rigid body MR-CT and MR-PET registration using MI and NMI, and found that the results for the two approaches were similar. In this particular case, the assertion that multi-resolution will improve registration is unfounded. This is due to the fact that any change towards a more uniform joint intensity histogram will increase the entropy (Subsection 3.4.2). Such changes may results from smoothing, and subsequently image blurring reduces the statistical power of joint intensity histogram for registration. Thus further experiments are needed to investigate the non-rigid registration using multi-level approaches with and without blurring for image block matching.

In these experiments, the multi-level sampling technique with a displacement median filter followed by Gaussian interpolation produced the best results. It outperformed the implementation that lacked a displacement field filter and one that maintained a fixed block size, as well as one that used a thin-plate spline interpolator. ADE decreased as the registration proceeded from coarse to fine level. With the Gaussian interpolator, the fast implementation produced similar results to the full implementation. It accelerated the processing by a factor of 6 without trading off registration accuracy, as measured by the average DE. In the experiments, 5 analytic forms of deformations were assumed as outcomes of a random process with an identifiable mean and PDF. One alternative approach to construct random distortions is by a suitably defined set of random parameter values for a single deformation, with a mean that corresponds to the case of no distortion, but it's harder to construct this for analytically independent deformations. It follows that the Gaussian δ width obtained in the Subsection 7.5.2 of a particular support may not necessarily be able to handle different analytical distortions. The choice of δ may have more to do with the extent (e.g. average displacement due to deformation) than the type of distortions and warrants further investigation.

Without subsampling, all similarity measures have comparable registration accuracy for both the intra- and inter-modality registration. When sparse sampling was used in registration, SCR produced superior results to MI in both the intra- and inter-modality simulations. A similar finding in global rigid registration was reported by Roche *et al.* [181]; they found that CR performed better than MI. The processing speed can be increased further with sparse sampling, particularly with [8,4,1] sampling as shown in Table 7.3. All similarity measures produced degradation in registration accuracy when sparse sampling was employed. However, SCR was consistently least affected and produced the lowest average DE. To understand better the discrepancy between MI and SCR, each similarity measure was computed as a function of translation for a 64×64 block in the proton density image with respect to the corresponding block in the T_2-weighted MR image (Figure 7.17).

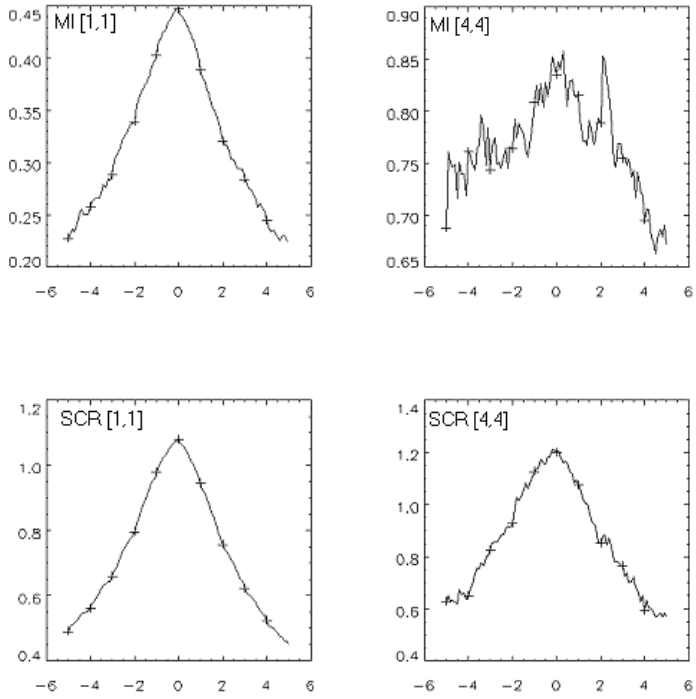


Figure 7.17: A comparison of SCR and MI as functions of translation (in pixels) along one axis. In this example, two MR image blocks (proton density and T_2-weighted) of size 64×64 were matched. The position of zero corresponds to correct registration. Notation [1,1] denotes full sampling, [4,4] denotes subsampling by a factor of 4 in each dimension together with 64-bin quantization.

The subsampling [4,4] was performed on the 64×64 image blocks without blurring. A surprising result was that, unlike SCR, MI manifested substantial "ripple" when sparse sampling was used. The ripple increases the likelihood that registration based on MI

will be trapped by a local maximum [219]. It helps to explain why SCR gave better performance than MI in Figure 7.13 when subsampling was used. It is worth noting that image block smoothing before subsampling does not necessarily improve performance. Any change towards a more uniform intensity distribution may perturb the joint intensity histogram, and hence affect the registration performance. A comparison of 3D non-rigid MR registration using multi-level subsampling of image block with and without smoothing is in Subsection 8.2.1.

In the current implementation, the regularization of the displacement vector field was performed after block alignment rather than by imposing an explicit constraint during local optimization. The median filter was more effective than the averaging filters in removing outliers in the sampled displacement vector field. The separate operation on the Cartesian components is a fast and efficient implementation of a median operator. In polar coordinates, due to the periodicity of orientation (phase angle), boxcar, Gaussian (not shown) and median filter (not shown) failed when operated separately on vector magnitude and orientation (Figure 7.18). On the other hand, the median filter, operating on Cartesian components, showed recovery of the displacement vector field better than the boxcar filter.

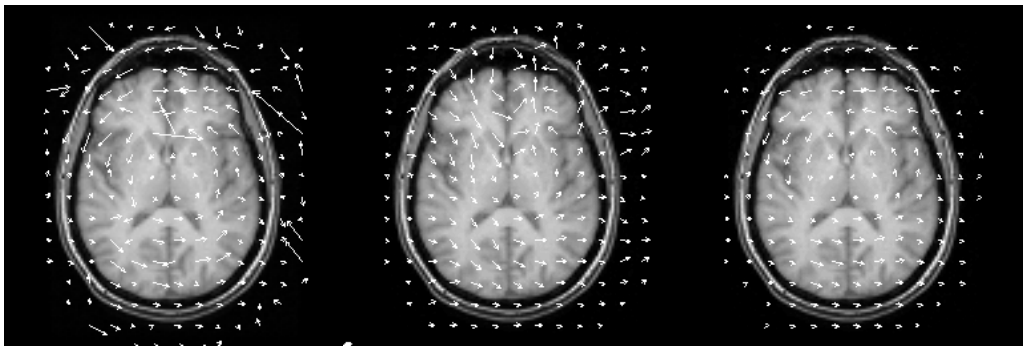


Figure 7.18: The displacement vector field obtained before and after filtering. The reference image was globally rotated. Measured displacement (left), displacement vectors after boxcar filter applied in polar coordinates (middle) and displacement vectors after median filter applied in Cartesian coordinates (right) are shown.

Cachier *et al.* [378] and Guimond *et al.* [379] also suggested the use of a Gaussian function to regularize the displacement field which, in their method, was derived from optical flow concepts. They employed a multi-resolution approach where coarse sampling of the displacement vector field was based on a pixel-to-pixel intensity difference. A fixed width Gaussian smoothing was applied to regularise the

displacement vector field obtained from image intensity gradient. In the proposed algorithm, a multi-level approach was also adopted. Unlike the method of Cachier and Guimond, the proposed coarse-to-fine implementation works directly on the displacement vector field which is derived from a block matching algorithm. In addition, the regularisation involves two processes: filtering and interpolation. A median filter is used to regularize the displacement vector field. It removes outliers in the vector field obtained from block matching. No explicit constraint is used in the local match. The Gaussian function, in this case, is a convolving kernel used to interpolate the displacement vector field from sparsely sampled vector data and the Gaussian width is adaptive to the sampling density.

7.7 Conclusion

A fully automatic non-rigid registration algorithm that uses a symmetric version of the correlation ratio as a matching criterion in a multi-level strategy has been developed and evaluated. The algorithm achieved improved registration for simulated intra- and inter-modality intra-subject examples. In each case, the average error between the recovered and the applied displacement decreased progressively with iteration using finer grid spacing. The symmetric correlation ratio in combination with regularization constraints imposed by a median filter and a Gaussian interpolator in the spatial domain appears to provide a flexible and robust automatic algorithm for non-rigid registration. The use of sparse sampling and the nearest neighbour Gaussian interpolation accelerate non-rigid registration towards clinically acceptable times.

Chapter 8

Validation of 3D non-rigid registration in multi-modality image

As described in the preceding chapter, studies were performed to optimize various strategies for 2D implementation. The results from these experiments show that the multi-level technique using a median filter followed by a Gaussian interpolation produces the best results. As an extension of the algorithm described in Chapter 7 into 3D, the aim of this chapter is to further evaluate the proposed algorithm for use in intra- and inter-modality non-rigid registration, for example, MR-MR, SPET-MR and SPET-CT in 3D. The image is divided into a set of overlapping 3D blocks positioned on a 3D grid. After the optimum transformation was found for each 3D image block, a 3D translational transformation was determined. The resulting displacement vector field, following interpolation using a Gaussian interpolator, fully defines the optimum transformation and provides non-rigid registration between a reference and a floating image. Accuracy of the technique was assessed using simulated and clinical data. In the rest of this chapter, the proposed algorithm is denoted as multi-level SCR (MSCR)

8.1 Method

The non-rigid registration problem is to determine a transformation T_u that deforms a floating image A to match the reference image B. The transformation T_u is parameterized by a displacement vector field u such that $A' = T_u A$. A detailed description of the algorithm, its formulation and implementation can be found in Chapter 7 (Section 7.2). Briefly, it can be described as follows:

- divide the reference and the floating images into overlapping image blocks,
- translate each floating image block with respect to the corresponding reference image block;
- find the translation that maximizes a similarity measure (SCR),
- from the displacement field form the optimum translation vectors,
- remove outliers
- interpolate the sparse sampled displacement vector field $u(x)$ to obtain the final transformation.

8.1.1 Image data

To evaluate the performance of the MSCR in 3D, four image data sets for intra- and inter-modality registration were used for the experiments. Three image data sets were used for the brain while one image data set was used for the thoracic-abdominal region. For the brain region, the data sets involved were: three intra-subject intra-modality T_1 MR image pairs for registration experiments plus ten inter-subject MR images for realistic distortions extraction, ten intra-subject inter-modality PD and T_2 MR image pairs and ten intra-subject inter-modality MR and SPET image pairs. For the non-brain region, an intra-subject inter-modality CT and SPET image pair was used.

The brain

The first (intra-modality) image set (Figure 8.1) consisted of three subjects. Two T_1-weighted MR images were acquired consecutively, with a few minutes pause for each of the three subjects. The original images comprised 256 slices of 0.976mm thickness measured center-to-center. Each slice consisted of 256×256 pixels of size 0.976×0.976 mm^2. No specific procedure was employed to ensure alignment between consecutive acquisitions. Each slice was reduced to 128×128 pixels (1.88^2 mm^2) and each data set was resampled to 128 slices with a slice thickness of 1.88 mm. From each data set, 99 slices were extracted for the intra-modality image registration experiments.

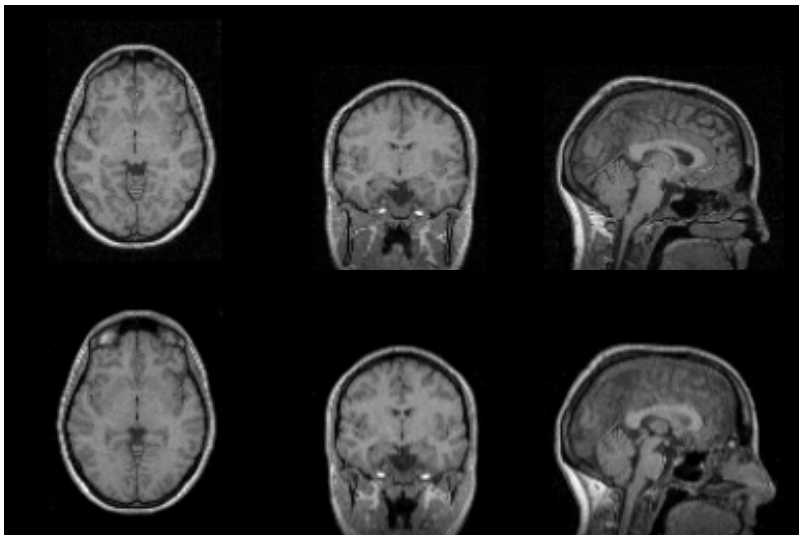


Figure 8.1: One of the three intra-modality T_1-weighted MR raw image pairs (128×128×99) is depicted in three orthogonal planes, from left: transverse, sagittal and coronal. The top and bottom images are of the same subject. Note that the two images are clearly misaligned.

In general, inter-subject registration accuracy in clinical data cannot be determined because the ground truth is not known. To circumvent this problem, inter-subject

deformation was used to simulate clinically realistic distortion between subjects. With this approach, it is assumed that inter-subject deformation encompasses intra-subject deformation in the sense that intra-subject deformations are expected to be smaller. Thus non-rigid registration between subjects will be a more difficult problem than the intra-subjects cases and a successful inter-subject algorithm is likely to succeed also in intra-subject registration.

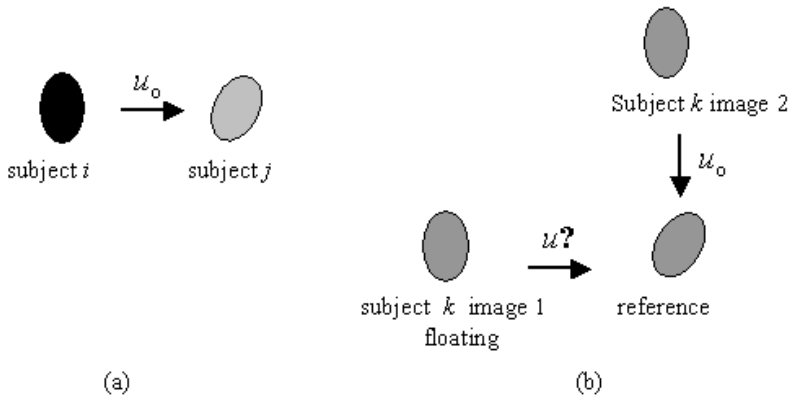


Figure 8.2: Illustration of the method used for validation. (a) Realistic displacement fields u_o are first determined by registering different subjects (i and j). (b) Of the two independently acquired images 1 and 2 for a subject k, one serves as the floating image whereas the other is deformed using the displacement field u_o to provide the reference image. Determining the displacement field u that best matches the floating image to the reference image tests the registration algorithm. Comparison is then made between the actual and derived displacement fields.

To simulate inter-subject distortions, realistic deformations were determined by matching T_1-weighted MR images between subjects. A version of the non-rigid algorithm (the only difference was the similarity measure) based on NMI, rather than SCR, was employed to reduce the possibility of bias. There were ten subjects (one image *per* subject) giving 45 possible pairings. Of these, ten pairs were selected randomly to provide a set of ten realistic displacement fields u_o. Then, image pairs were chosen from one of the image data sets. All image pairs were first registered using global rigid registration. One of the two images was then deformed using each of the ten displacement fields u_o obtained previously. The deformed image provided the reference image B, the second image acting as the floating image A (Figure 8.2). The MSCR non-rigid registration algorithm was then applied to determine the displacement field u that best aligns the floating image to the reference image. Finally, the recovered displacement field u was directly compared with the applied deformation u_o to compute registration accuracy as outlined in Subsection 8.1.3. Two further experiments were also conducted to study the effect of fixed block size and effect of image blurring to a lower

scale on registration accuracy, with default implementation. In the first case, the results of registration accuracy using fixed block size [32^3,32^3,32^3] were compared with the adaptive image block size implementation. In the second case, the reference and floating MR images were smoothed prior to registration. After smoothing, each image was subsampled to generate coarse images in resolution for registration, and the results were compared with the registration without blurring. Since value 5 was closest to the subsampling factor 4, a 5×5×5 boxcar filter was used for blurring the images to a lower scale.

The second (inter-modality inter-subject) data set comprised pairs of T_2-weighted and proton density (PD) MR images, one pair for each of the ten subjects. These image pairs were used to evaluate the accuracy of registration of high resolution inter-modality images using the MSCR. The T_2-weighted MR image was acquired immediately after acquisition of the PD MR and was, therefore, in near-perfect registration with the PD MR. This alignment was verified to some extent by the rigid registration, which found the optimum transformation parameters in the vicinity of zero translations and rotations for each pair of images. Each image consisted of 19 slices of 128×128 matrix with a pixel size of 1.80×1.80 mm^2 and a slice thickness of 7 mm. All MR images were resized to give a cubic voxel of side 1.80 mm and data were padded with zeros to give a total of 99 slices (Figure 8.3).

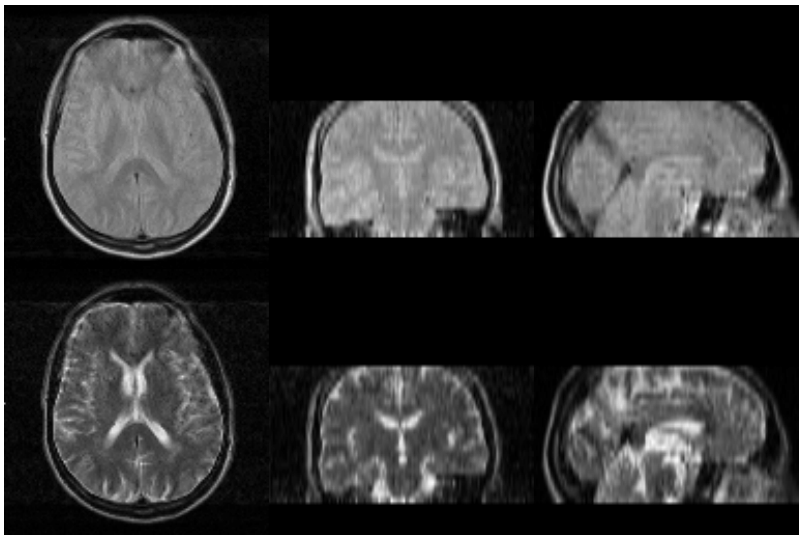


Figure 8.3: One of the ten inter-modality images pairs, comprising a proton density image (top) and T_2-weighted image (bottom) of the same subject, is shown in three orthogonal planes. The images consisted of 19 slices, each with 128×128 pixels (1.80×1.80×7.0 mm^3 voxels) when acquired but downsampled to 128×128×99 matrix with cubic voxel size of 1.80 mm. Note that the images extend nearly, but not quite, to the top of the skull.

The third image data set was used for inter-modality registration. It consisted of pairs of MR and SPET images from Subsection 4.6.1 (Figure 8.4), one pair for each of the ten subjects. By contrast to the second data set of high spatial resolution MR images, these image pairs assess the performance of the MSCR for images of very different spatial resolution. The matrix sizes were 128×128 pixels (1.88×1.88 mm^2) of 124 slices (1.5 mm) in MR, and 64×64×48 (3.59×3.59×3.59 mm^3) in SPET images. The MR and SPET images were resampled to give cubic voxels (1.88^3 mm^3) and of matrix dimension 128×128×99. Global rigid registration was performed on each pair of SPET and MR images before the non-rigid registration experiments.

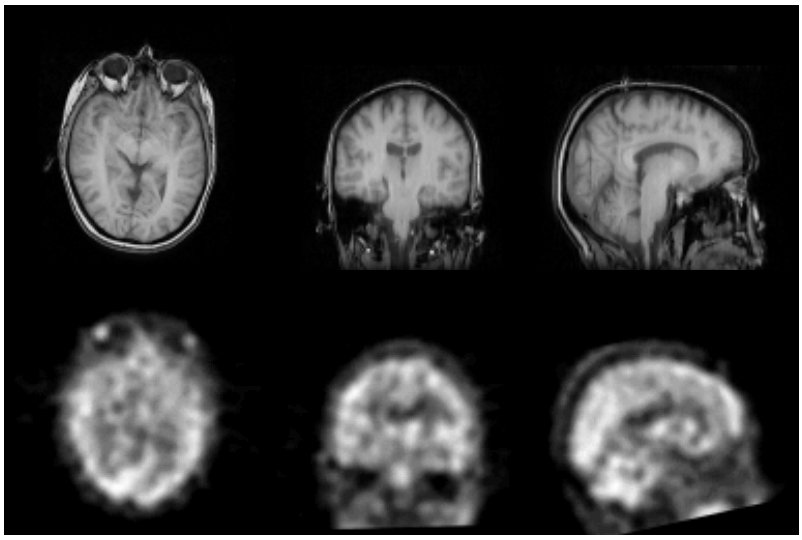


Figure 8.4: An image pair from the third data set. An MR image (top) and a corresponding SPET image (bottom), resampled to a common matrix comprising 128×128×99 cubic voxels of size 1.875 mm^3, are shown (left to right) in transverse, sagittal and coronal planes. The images are shown following a global rigid pre-registration.

The torso

To investigate the feasibility of the MSCR to register SPET and CT in the non-brain region, Zubal CT and simulated SPET were used. The CT image comprised 128×128×112 cubic voxels of size 4mm. A SPET image was simulated based on the segmented Zubal CT image (Chapter 6, Section 6.2). The source distribution assigned to the Zubal phantom mimicked the ^{67}Ga-citrate distribution. It was forward projected in a 64×64 grid of 8×8 mm^2 pixels with 64 views over 360 degrees. The resulting projections incorporated non-uniform attenuation, depth dependent resolution and Poisson noise. In order to avoid streak artifacts in FBP and its poorer noise characteristics (Chapter 5, Subsection 5.4.7), the 3D projection data were reconstructed with OSEM (one iteration of four subsets 1OS4) to produce a 64×64 array with 8 mm

slice thickness. The entire SPET volume was distorted non-linearly (Figure 8.5). A three dimensional global thin-plate spline [55,198,379] with 46 control points was used to warp the thoracic-abdominal region. Of these, 29 control points were placed around the thoracic region and body boundaries to maintain spatial continuity and topology of body tissues. The remaining 17 control points were placed at the outer surface of the liver. Displacements selected at random from a uniform distribution ranging from zero to ±5 voxels were applied independently along the x-, y- and z-axes. Five independent data sets were created. To assist in the selection of control points, a split window display showed intersecting axial, coronal and sagittal slices. A mouse-controlled cursor in the form of crosshairs indicated the position of the intersection point. The coordinates of every voxel in the entire CT volume were transformed with the aid of trilinear interpolation. The transformed CT image was taken as the reference image. Before registration, the SPET (floating image) was resized from 64×64×56 array to the 128×128×112 format of the CT volume to facilitate the non-rigid registration.

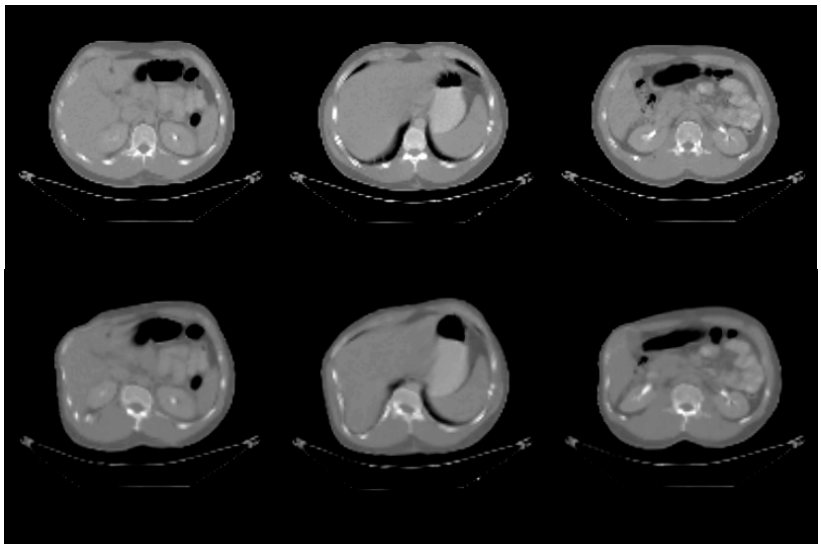


Figure 8.5: Selected transverse slices of the Zubal CT image before (top) and after (bottom) 3D thin-plate warping. Note the different appearance of image after the non-linearly and spatially varied distortion.

8.1.2 Implementation

The algorithm of chapter 7 (section 7.3) was employed, incorporating a multi-level approach, to determine the displacement vector field. The non-rigid registration procedure started with local block matching using individual image blocks. The registration was performed using translation in 3D and the quality of the fit was determined by a similarity measure. Interpolation of the displacement vectors between grid nodes was implemented as a convolution (see Equation 7.7) of the sampled

displacement vector field with a Gaussian kernel. The displacement vector field $u^{(\kappa)}(x_i)$ at position vector x_i at level κ was obtained by interpolation,

$$u^{(\kappa)}(x_i) = \begin{cases} u^{(\kappa-1)}(x_i) + \phi^{(\kappa)}(x_i) * \tilde{v}^{(\kappa)}(x_i), & \kappa > 0 \\ 0, & \text{otherwise} \end{cases},$$

where the Gaussian kernel at iteration κ is denoted by $\phi^{(\kappa)}(x_i)$, and $\tilde{v}(x_j)$ represents a residual displacement vector sampled at a grid node voxel *j*. It is important to state that every image block was registered using its own voxel-based information content. The sampled displacement vectors were median filtered over a 3×3×3 neighbourhood in a postprocessing step. Median filtering suppressed outliers and smoothed the displacement field. The displacement was updated by adding the current residual displacement field to the previous estimate $u^{(\kappa-1)}(x_i)$. The only restriction imposed on the magnitude of the displacement was that no displacement greater or smaller than the median of the neighbourhood was allowed. That is, all displacements were replaced by local medians.

At each iteration, the optimizer (the standard simplex algorithm [127]) minimised a negated joint-intensity based similarity measure (SCR, MI or NMI). To determine the best spatial correspondence between a pair of image blocks, iterations of the optimization algorithm are repeated until either a maximum number of iterations were reached or the similarity measure changes by less than a chosen criterion. The registration is progressively refined using the strategy described in Subsection 7.3.1. To speed up the registration process, a downsampling scheme [4,2,1] was employed. Experiments in [258] showed that this downsampling scheme worked best with intensity rebinning [64,48,32]. For example, the 64^3 image block was rebinned from 256 intensity bins to 64 bins, and sampled every fourth voxel along the x-, y- and z-axes, thus generating 16^3 samples for the computation of the similarity measure. The computed displacement vector field was used to map the floating image onto the reference frame with the aid of the trilinear interpolation. Unless stated otherwise, the use of the simplex optimisation, median filter, downsampling schemes [4,2,1], intensity rebinning [64,48,32] and the trilinear intensity interpolation is referred to as the default implementation in the rest of this chapter.

8.1.3 Evaluation of 3D non-rigid registration

To assess the registration accuracy of the MSCR in registering intra- and inter-modality images, voxel error was computed using $\varepsilon_i = \|u_i - \tilde{u}_i\|$ for each voxel (see Subsection 7.3.4). It measures the discrepancy between the recovered displacement $\tilde{u}_i$ with respect to the ground truth (i.e. applied distortion u_i) by the registration process at each voxel. Summed over all voxels in the image, ε gives the RMS displacement error (DE) by $\mathrm{DE} = \frac{1}{\sqrt{N}} \sqrt{\sum_{i=0}^{N-1} \varepsilon_i^{\,2}}$, where N is the image size. Average DE over all images was also computed for each intra- and inter-modality registration using the default implementation (Subsection 8.1.2). Three downsampling schemes were used: [1,1,1], [4,2,1], [8,4,1] and [16,8,2]. In the PD-T_2 registration experiment, volume padded above and below image volume data (see Figure 8.3) was excluded from the DE computation because of the lack of tissue content and N was correspondingly reduced. For SPET-CT registration in the thoracic-abdominal region, N was taken over all non-background voxels in the CT volume. Since in the 2D study, SCR showed a consistent performance with slightly better registration than CR1 and CR2 (Subsection 7.5.4 and 7.5.5), all results reported in this chapter were compiled for SCR, MI and NMI. The difference between two registrations was assessed statistically by a paired comparison of DE over all image pairs using the two-tailed paired t test. Histograms of voxel error ε before registration and after the first, second and third iteration for each intra- and inter-modality registration was also computed over all corresponding studies. Thus, for each registration, four histograms of ε are combined, showing the distribution of voxel error with respect to the ground truth using MI, NMI and SCR.

8.2 Results

8.2.1 Intra-modality registration

The results presented here were obtained from ten randomly selected volume pairs. A typical frequency distribution of the voxel error using SCR, NMI and MI for the T_1-weighted MR intra-modality registration is illustrated in Figure 8.6. The result was pooled over all the ten studies to produce histograms of voxel error before registration and after the first, second and third iteration. The histogram shows voxel error distribution for the misregistration with respect to the applied distortion (i.e. ground

truth) using SCR and NMI, with [4,2,1] sampling scheme and the default implementation. To further illustrate the difference of performance between SCR and MI, histograms of voxel error using [16,8,2] sampling scheme were also given.

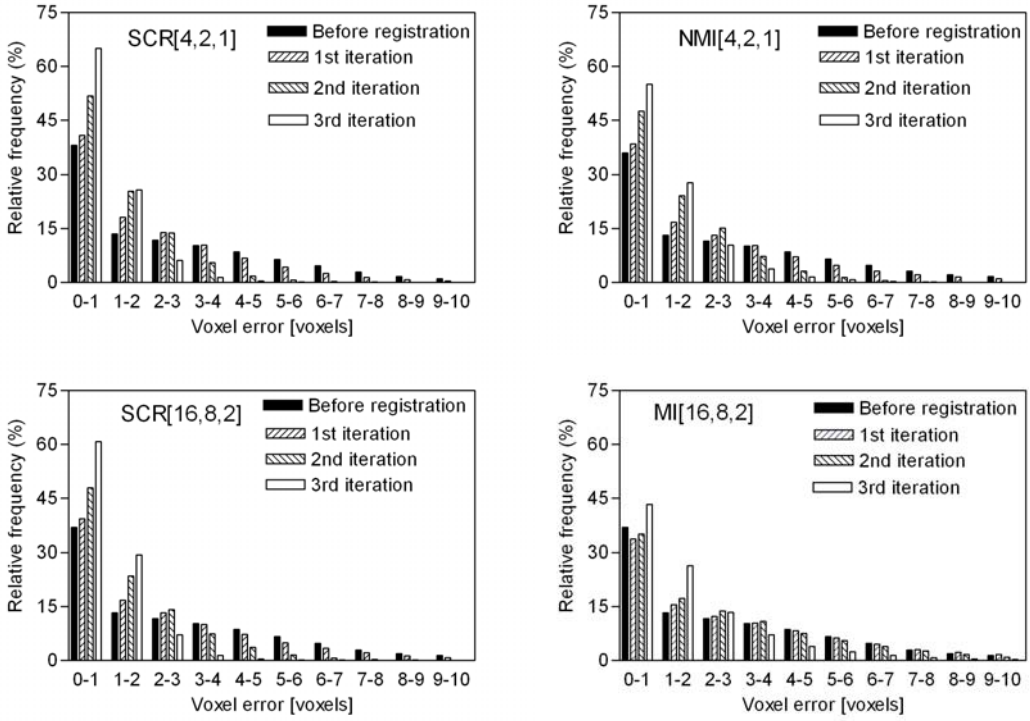


Figure 8.6: Histogram of voxel errors over all image voxels pooled over ten studies, before registration and after one, two and three iterations. Default implementation was used for the registration. Downsampling scheme [4,2,1] with SCR (top left) and NMI (top right), and [16,8,2] with SCR (bottom left) and MI (bottom right) are shown.

The histograms show a gradual improvement of registration with iteration number: the voxel error diminishes and more image voxels are registered within an accuracy ±1 voxel. The proportion of registrations with accuracy ≤ 2 voxels increases from 59% after the first, 77% after the second, and 92% after the third iteration. Similar behaviour was observed in the voxel error histogram for MI and NMI. Since deformations were obtained with NMI, its voxel error histogram provides a consistency test and, in this sense, acts as a reference (see Figure 8.6 top right). However, a difference in performance was observed between SCR and MI when sparse sampling [16,8,2] was used. For MI, the proportion of voxels with 0-2 voxel accuracy was steady until the final iteration (50.3%, 49.2%, 52.3% and 69.7%) but SCR showed consistent increase with iterations: 50.3%, 56.1%, 71.6% and 90.2%. With the sparse sampling, MI and SCR produced a 92% of registration with accuracy ≤ 2 voxels over all non-background voxels with [4,2,1] but SCR started to incur progressively greater success rate than MI

and NMI in [8,4,1] and [16,8,2]. The proportion (i.e. proportion of registrations with accuracy ≤ 2 voxels) for SCR was independent of the sampling scheme. For MI, the proportion dropped by about 20% in the sparsest sampling scheme. For NMI, the proportion also dropped from 92% with [4,2,1] and [8,4,1] to 80% with [16,8,2]. The results suggest that SCR is more robust than MI and NMI for non-rigid registration when sparse sampling is employed. The computational cost advantage of sparse sampling is substantial. In completing three consecutive levels of block matching, the most aggressive downsampling [16,8,2] demonstrated a fivefold increase in speed over [1,1,1], whereas [4,2,1] and [8,4,1] showed threefold increase (Table 8.1). Examples of images before and after registration are shown in the checkerboard pattern (Figure 8.7). Following registration, the two images appeared to be significantly well aligned. There is a reduction of visible mismatch following registration. The good alignment of the brain surface and other features demonstrates qualitatively the accuracy of the registration method. It's interesting to note that the reference image appears anatomically impossible. It indicates that although inter-subject transformations are "realistic", application of these transformations to a real image does not necessarily produce "realistic" images.

	[1,1,1]	[4,2,1]	[8,4,1]	[16,8,2]
Average time (min)	29.1	10.8	10.1	5.86
Relative time	100%	37%	34%	20%

Table 8.1: Average processing time for 3D registration with SCR for different sampling schemes, as measured on 1.7 GHz Pentium PC. Note that a sampling factor of two was used at the third level in the [16,8,2] sampling scheme but not in the other three sampling schemes, which may partly account for the 80% reduction in time.

To further compare the performance of SCR to MI and NMI for T_1-weighted MR intra-modality registration, displacement error (mm) over all image voxels of ten studies for various down-sampling schemes are shown (Figure 8.8). All three measures gave a similar error distribution and demonstrated comparable accuracy in the mean DE at [1,1,1] and [4,2,1]. For example, with [1,1,1] sampling scheme, the mean (±SD mm) DE for MI, NMI and SCR are 1.6±0.4, 1.6±0.4 and 1.7±0.3 respectively. Similar mean (±SD mm) DE was found for MI, NMI and SCR using [4,2,1] sampling scheme. The performances of the SCR and MI became statistically differentiable when the sparsest sampling scheme [16,8,2] was used. In the 2D registration study performed in Chapter 7 (Figure 7.13), differentiation emerged after [4,2,1], possibly reflecting fewer samples available in 2D to compute the joint intensity histogram. In the 3D scheme [8,4,1], the

mean (±SD mm) DE were gradually reduced from 1.8±0.4, 1.6±0.3 and 1.7±0.3 corresponding to MI, NMI and SCR respectively. In the most aggressive down-sampling scheme of [16,8,2], degradation is observed although it is not uniform across the similarity measures. MI suffered most (P<0.05) and NMI and SCR were least affected. In this case, the mean (±SD mm) DE for MI, NMI and SCR were 3.2±0.6, 2.4±0.6 and 1.8±0.3, respectively.

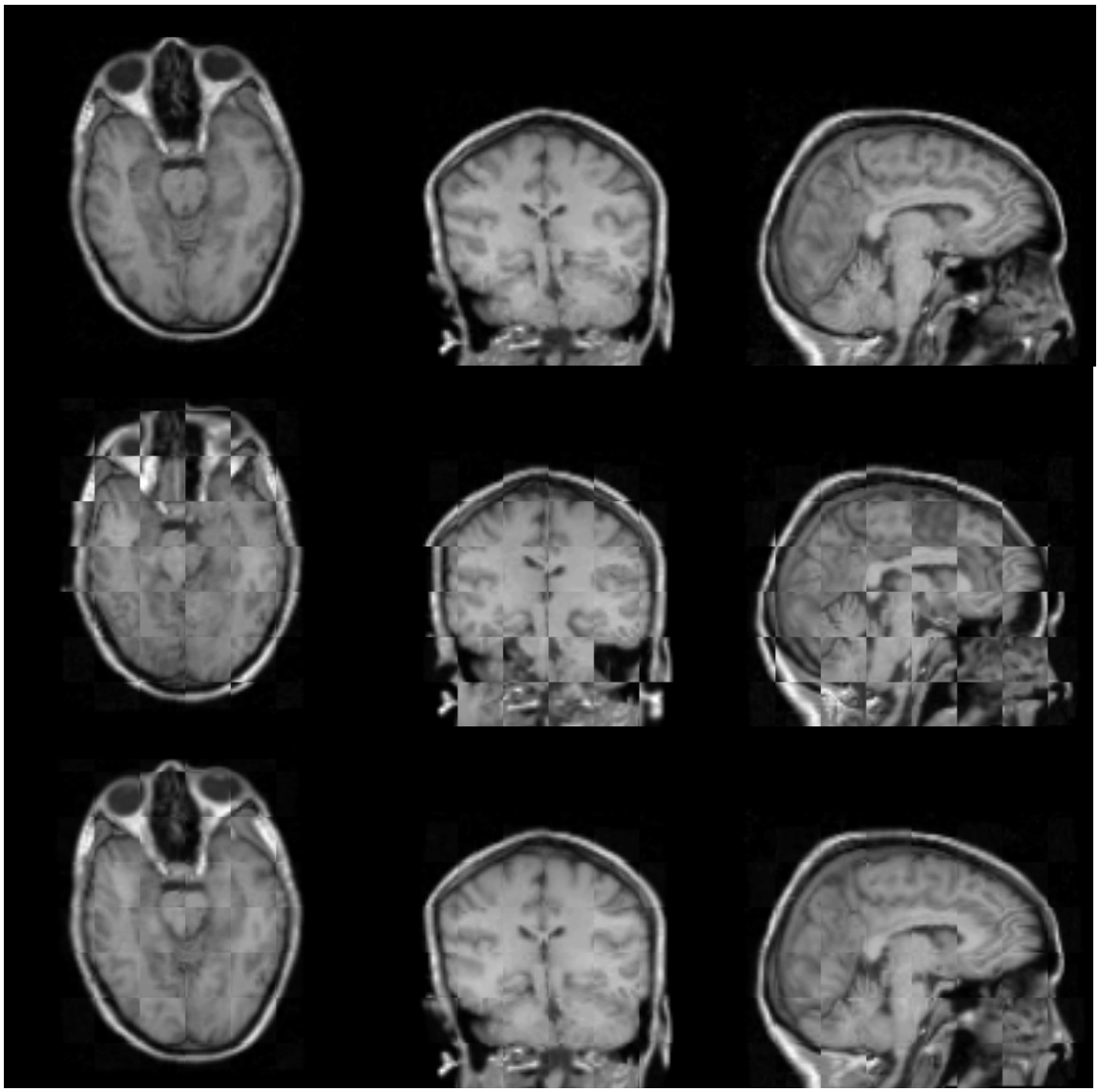


Figure 8.7: The results of intra-modality 3D registration shown using the checkerboard display. The columns, from left to right, correspond to the transverse, coronal and sagittal sections. Shown are the reference image (top row), the reference image interlaced with the floating image prior to (middle row) and after the registration (bottom row).

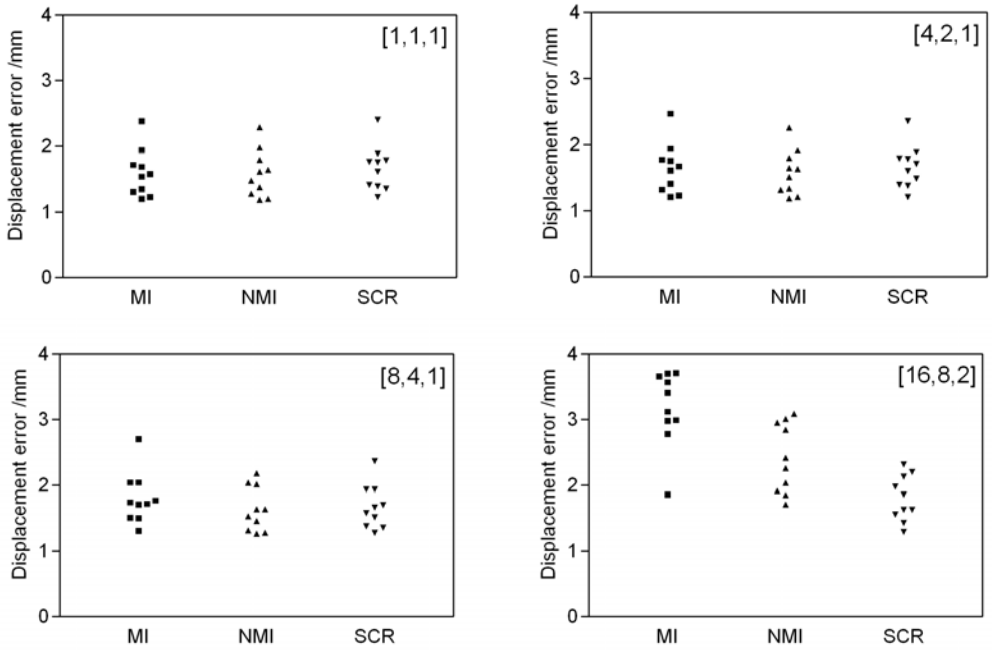


Figure 8.8: Comparison of MI, NMI and SCR in 3D intra-modality simulation experiments. Scatter plots of DE is shown for several downsampling schemes. Note that the scale is the same for all plots.

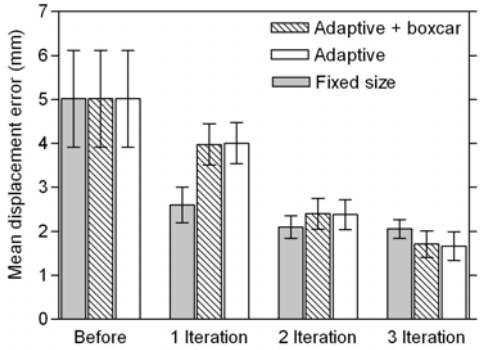


Figure 8.9: Mean displacement error (mm) for intra-modality T_1 MR registration before registration and after various iterations, using default implementation and SCR. DE were pooled over voxel error over all image voxels of 10 independent image pairs. Error bar refers to the standard deviation over ten studies. 'Fixed size' denotes registration using $[32^3,32^3,32^3]$, 'adaptive + boxcar' denotes registration using $[64^3,32^3,16^3]$ with blurring and 'adaptive' refers to default implementation.

The mean displacement error against iterations for default implementation (adaptive block size) with and without boxcar filtering, and fixed block size $[32^3,32^3,32^3]$ is given in Figure 8.9. The results show that registration accuracy was improved irrespectively to the three implementations but the improvement was not uniform across all implementations. The adaptive block sizes improved registration accuracy progressively with iterations. For fixed block size, the improvement flattened out at the second and third iteration although a relatively large improvement comparing to the adaptive block sizes was produced after the first iteration. The results shows that adaptive block size

was more effective than fixed block size for registration. This was confirmed by statistical tests (two-tailed paired t-test, $p<0.05$) that there was a significant difference of mean DE between fixed size and adaptive sizes. However, there was no statistical difference between mean DE with and without boxcar filtering for default implementation. The mean (±SD mm) DE for fixed size iterations, adaptive size iterations with and without blurring to a lower scale were 2.1±0.2, 1.7±0.3 and 1.7±0.3 respectively.

8.2.2 Inter-modality registration

The purpose of this study was to assess the MSCR for 3D inter-modality non-rigid image registration. Brain images of comparable high resolution (proton density and T_2-weighted MR) were registered as well as images of different resolution for MR-SPET registration. Torso images obtained with CT and simulated SPET were also registered.

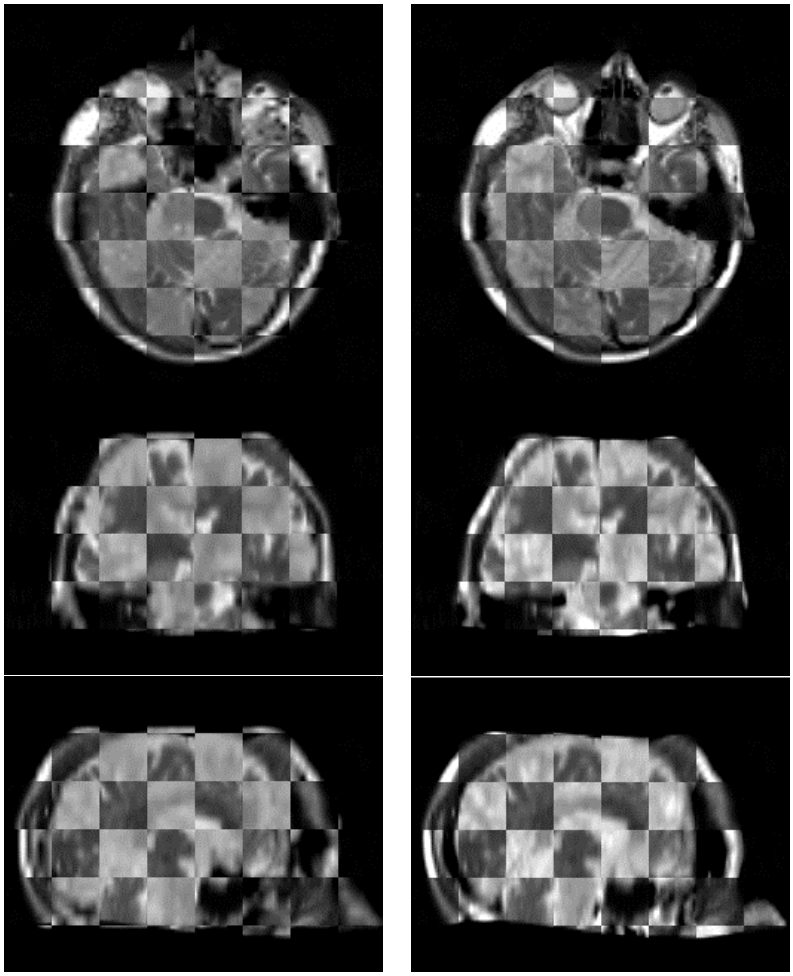


Figure 8.10: The results of inter-modality for proton density and T_2-weighted MR 3D non-rigid registration using SCR and default implementation. The columns, from top to bottom, correspond to the transverse, coronal and sagittal sections. The checkerboard display fuses the reference image and the floating image before (left) and after (right) registration.

Proton-density and T_2-weighted MR images of the brain

Ten independent image pairs of proton-density and T_2-weighted MR images were used. A typical image pair is shown before and after non-rigid registration in Figure 8.10. The fused images are illustrated using checkerboard display. This enables visual determination of the accuracy of boundary alignment in various regions. Even a subtle shift of an edge can be seen clearly. Before registration, mismatches between the two images can be seen, for example, in the forehead region and eyeballs. After registration, matching of anatomical structures is apparent between the images. In the registered floating image, the alignment of the skull (both inner and outer surface) is excellent and there is a good alignment of the brain surface throughout the image volume. To evaluate quantitatively the performance of the MSCR in inter-modality registration, a histogram of voxel errors was computed before registration, and after each iteration. Figure 8.11(top left) shows the histogram of voxel error pooled from the results of all ten PD-T_2 MR volume pair registrations using SCR and the default setting before and after each iteration.

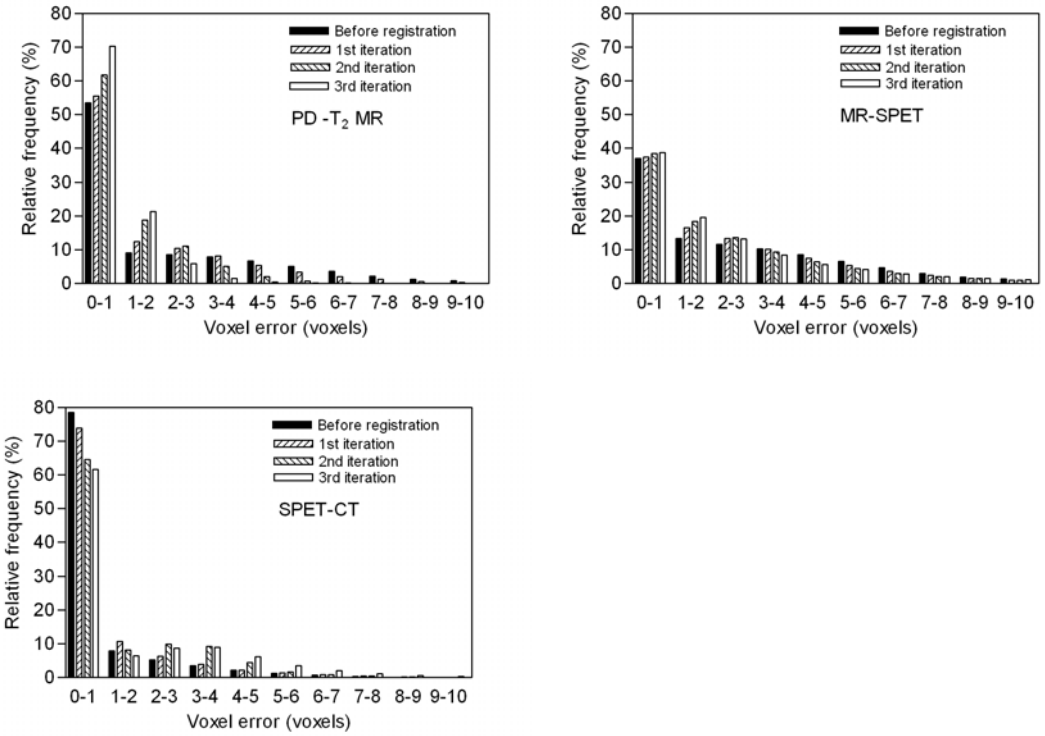


Figure 8.11: Histograms of voxel error pooled over all studies, before registration and after each iteration of the inter-modality registration algorithm. Default implementation, with sampling scheme [4,2,1] and SCR were used. Histograms are constructed from results of registration of (top left) proton density (PD) and T_2-weighted MR images of the brain, (top right) T_1-weighted MR and SPET images of the brain, and (bottom left) CT and simulated SPET images of the torso. The error is expressed in units of MR or CT voxel size, as appropriate. The histograms are aggregated from ten image pairs in brain studies and 5 image pairs in the torso. Voxels outside the body were excluded.

The proportion of registrations with accuracy ≤ 2 voxels improved progressively from 63% (before registration) to 68%, 81% and 92% after each successive iteration of the default implementation with the SCR similarity measure. The other measures (MI and NMI) showed similar results. To further illustrate the results, summary statistics was also computed for all the registrations. For example, the mean (±SD mm) displacement error for MI, NMI and SCR were 2.1±0.3, 1.9±0.3 and 2.0±0.3 respectively, with the default sampling scheme. No statistical difference (p<0.05) was found among the three measures. Uniformity of performance is also demonstrated by the proportion of registrations with accuracy ≤ 2 voxels; (92±4)% for each measure and confirmed by visual inspection. The performance of MI, NMI and SCR was virtually identical, and a paired two-tailed t test indicated no significant difference at p<0.05 level. Again, the statistical measures reflect the change in visual quality of registration demonstrated in the checkerboard display (Figure 8.10). As expected, registration reduces the voxel error.

MR and SPET images of the brain

In the MR-SPET non-rigid registration experiment, a total of 10 independent image pairs were employed. All image pairs were first globally registered with NMI. Distortions were applied to each reference (SPET) image. The distorted reference image and the floating (MR) image were registered with the default settings (Subsection 8.1.2). Based on the recovered displacement field, voxel error and displacement error were computed with respect to the ground true (i.e. the applied displacement field) as described in Subsection 8.1.3. The results show that there was little, if any, improvement of mean DE with the default implementation regardless of the three similarity measures: MI, NMI and SCR, and was confirmed statistically (paired two-tailed t-test, p = 0.05). The mean (±SD mm) DE before registration was 5.0±1.1 and after registration was 5.0±1.6, 4.7±1.4 and 4.5±1.1 for MI, NMI and SCR respectively. The results suggest that non-rigid registration of the MR and SPET images was not effective. The MSCR appeared to produce little effect on improving the mean DE and registration accuracy (≤ 2 MR voxels) regardless of the choice of the similarity measure.

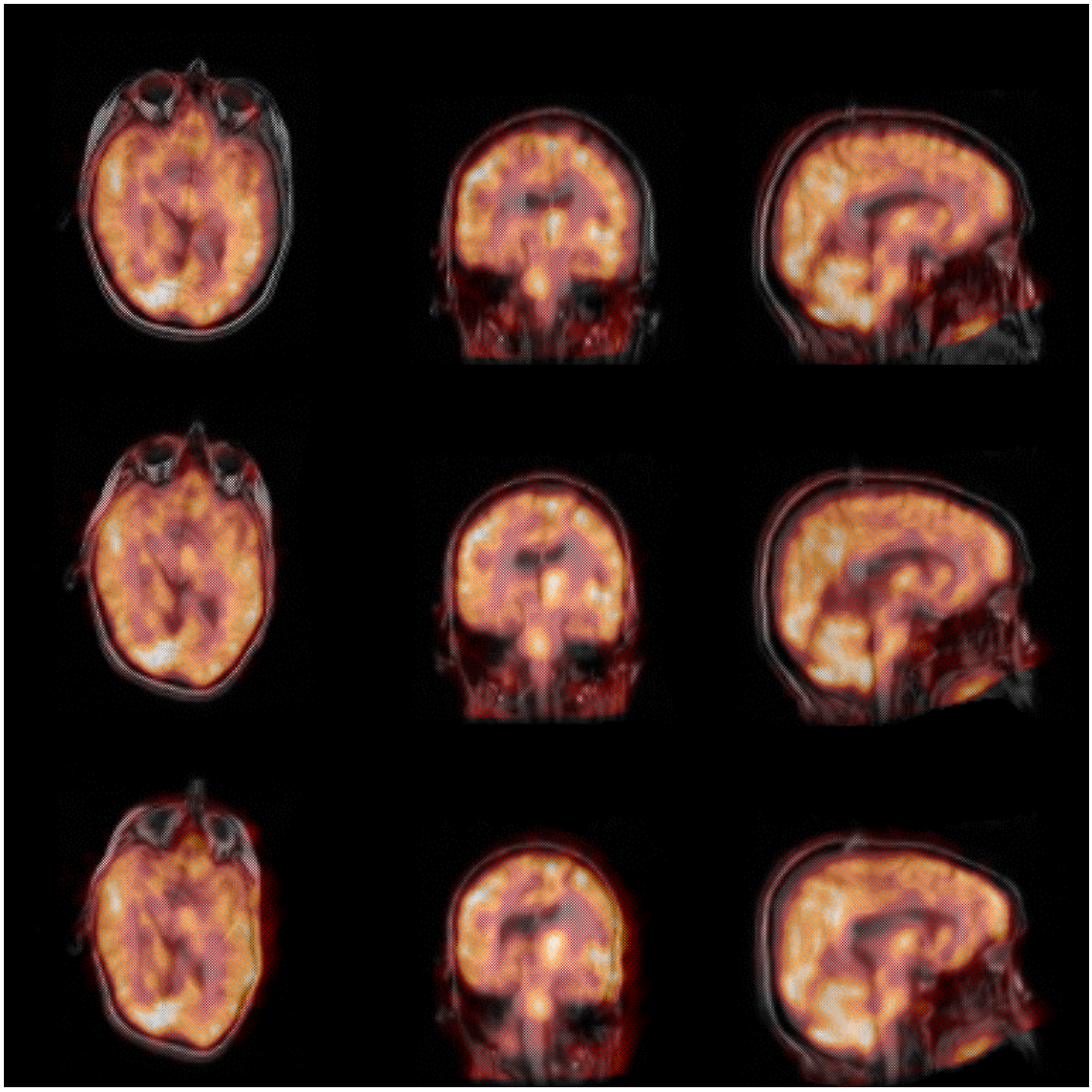


Figure 8.12: The results of SPET-MR 3D registration using default implementation. Shown from left to right, are the transverse, coronal and sagittal sections. The SPET and the MR images are first globally registered using a rigid-body transformation and the transformation applied to the SPET image. Reference image (SPET), rendered in 'red temperature' scale, overlays a grey scale image representing (top) the floating MR image before registration, (middle) the non-rigidly registered floating image, and (bottom) the correctly transformed floating image. Note that it is difficult to distinguish between these two cases.

A typical histogram of displacement errors (Figure 8.11 top right) shows that registration improved marginally after three iterations and in all three cases (MI, NMI and SCR) only a 10% improvement was found. The plot also reveals that large voxel misregistrations (error ≥ 4 MR voxels) persisted unlike the PD-T_2 MR registration. To account for the poorer resolution in SPET, it is more appropriate to express the histogram in units of the SPET voxel size of 3.59 mm. Again, the improvement of image alignment was weak, only 8%, compared to 42% in intra-modality MR registration, and 30% in inter-modality PD-T_2 MR registration. The proportion of

voxels with registration error ≤ 1 voxel of SPET was 50.3% (before registration), 54% (1 iteration), 56.9% (2 iterations) and 58.5% (3 iterations). Similar results were found for MI and NMI, and no significant ($p < 0.05$) difference was found among the three measures. A typical SPET-MR brain images before (top) and after (middle) registration is shown in Figure 8.12. In order to evaluate the registration visually, a colour overlay display is used. The reference image (SPET) is rendered in the "red temperature scale" and it overlays the registered floating image (MR) in gray scale. To visualize differences of the recovered distortion and the applied distortion, an overlay of the reference image over the correctly transformed floating image is also shown (bottom, Figure 8.12). Because of the low resolution in SPET and the absence of many anatomical structures in the SPET image, it is difficult to distinguish between these two cases.

CT and SPET images in non-brain region

In this experiment, the Zubal CT image was distorted using a non-linear transformation (thin-plate spline) and taken as the reference image while simulated ^{67}Ga SPET was the floating image. The thin-plate spline produces a spatially varying non-affine distortion in the entire volume. The MSCR was applied to the image pair to recover the applied distortion. An overlay display of volumetric image pairs for SPET and CT in the transverse direction is shown in Figure 8.13. The registration algorithm did not substantially improve visual alignment. The overlay images show misalignment of liver despite a good registration visually in other parts of the body. For example, the surface of the liver is aligned erroneously with the boundary of the body (right image of Figure 8.13) but the spine appears well aligned.

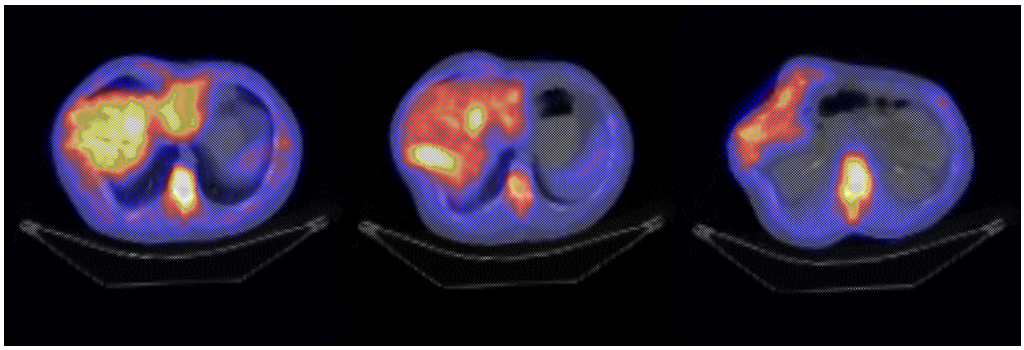


Figure 8.13: The floating image (SPET) in "STD Gamma-II" colour scale overlays the reference image (CT) in gray after non-rigid registration. The mismatch of the liver image between SPET and CT in three axial slices from upper (left), middle and lower (right) abdominal region is not improved after registration.

Since the global spline deformed the entire image domain, it would also deform the parts of the image occupied by background voxels. Background distortion is meaningless except near the body boundary where the displacement of background voxels represents expansion of the body boundary. Background voxels are also likely to produce spurious registration errors. For these reasons, background voxels that are not adjacent (in the sense of belonging to the same image block) to the body boundary are excluded from the computation of the displacement error. A quantitative assessment of registration performance is given in Figure 8.11 (bottom left), which plots the histogram of the voxel error computed over five independent distortions applied to a CT and a simulated ^{67}Ga SPET image pair. The histogram shows that the proportion of voxels with registration accuracy ≤ 1 voxel of SPET is progressively deteriorated from 86.5% (before registration) to 84.7%, 72.8% and 68.1% corresponding to each successive iteration. The result indicates clearly that registration did nothing to improve misalignment between the CT and SPET image pair. The results also reveal that the relative prevalence of displacement error ≥ 2 voxels increased as the registration progressed from second iteration to the third iteration. The incidence of large errors (≥ 4 CT voxels), which is more likely to be recovered using a smaller image block, rises from 5% after the first iteration, 7.9% after the second iteration and finally 14% after the third iteration. Poor performance was also found with MI and NMI similarity measures. No significant difference (p<0.05) was found among the similarity measures.

8.3 Discussion

In this chapter, an automatic method for non-rigid registration has been described and a 3D implementation has been demonstrated. Validation is an important step to demonstrate the quality of the MSCR registration technique in terms of accuracy, precision, robustness and reliability. A more detailed discussion of the issues of validation for image registration is found in Section 2.5. The main issue in measuring accuracy is that ground truth is not available for clinical data. Therefore, simulations of clinical images and deformations were used. Simulated data permit controlled evaluation over a wide range of conditions (e.g., image modality, distortion, or noise levels). The performance of the proposed registration was evaluated with 3D images simulating inter-subject distortions for intra-modality (T_1-weighted MR) and inter-modality studies (PD-T_2 MR images and MR-SPET images) in the brain region. Based

on the same approach, performance in the non-brain region was also assessed by comparing the known and the measured displacement following 3D inter-modality (SPET-CT) image registration. Comparison of registration was also performed to test the effectiveness of SCR, MI and NMI. The assessment of performance of the non-rigid registration is, in general, an open issue. A number of approaches were used to assess the algorithm, providing insight into the performance and the properties of the different similarity measures. First, for visual assessment, a checkerboard or a colour overlay display was used. It is simple, fast, and intuitive. For registration involving high-resolution image pairs, subtle differences along edges and small shifts at organ surfaces were easily detected, as shown in Figure 8.7 and Figure 8.10. Even for images of different resolution, the fused display offers a quick evaluation of any visible mismatch (Figure 8.12 and Figure 8.13). Second, registration accuracy was quantitatively assessed by computing the displacement error DE, relative to known distortions, over all voxels in a 3D image and the averaged values of DE over multiple images. For a better understanding of the effectiveness of the registration, a histogram of the voxel error was obtained to determine how the distribution of accuracy changes as registration progresses with successive iterations. Finally, the statistical significance of the difference between the means of DE distribution over similarity measures was determined using standard statistical tests.

In the present study, it is also noticed that 92% of voxels have a voxel error (≤ 2 voxels of MR) over all image voxels of the ten multiple studies for the intra-modality MR registration. This may not be good enough for clinical practice as a reliability level of at least 99% may be required. This may be due to the fact that the applied distortion deformed the entire image domain, including background and non-background voxels. Background distortion is meaningless and likely to produce erroneous registration errors. A statistical test was applied to determine whether there was a significant difference between mean DE with and without background. A significant difference of about 5% DE was found (paired two-tailed t-test, $p<0.05$) when DE was confined to within the body boundary and background voxels adjacent to the body boundary. In the case of non-rigid registration, the displacement error comprises uncorrected globally linear and non-linear distortions. The local algorithm should readily pick up any globally linear component of misalignment. However, it is common practice to first perform affine registration. The motivation is to bring the images into sufficiently close alignment to

allow local misalignments to be captured without too many spurious results. Scaling distortions between images can be accounted for by incorporating image voxel sizes in transformations, and shear deformation is unlikely to occur in the medical image, therefore rigid body registration is often used to remove global misregistration. As a first order improvement in the image registration, the global linear misregistration is accounted by a global rigid registration. The local misalignment is a second order misregistration that is removed by local translation. In aggregate, this amounts to globally nonlinear transformation. In this chapter, non-rigid registration is the main focus. Validation of the global rigid registration for images of the brain and torso can be found in Chapter 4 and Chapter 6, respectively.

Based on the experiments with MR images, results indicate that the MSCR non-rigid matching is capable of reducing local misregistration, irrespective of the choice of the similarity measure (SCR, NMI or MI). The improvement of registration with adaptive block sizes is probably due to the fact that registering larger blocks can reduce larger scale distortions and thus reduce errors which vary slowly across the image [373]. Smaller blocks have the capability to capture small-scale misalignment, perhaps those associated with smaller image structures, allowing a smaller displacement error (i.e. higher spatial frequency error) and thus a more accurate alignment. An application of particular interest for the MSCR approach is inter-modality registration. The MSCR was applied to MR-SPET, CT-MR and MR-MR (T_2-PD) non-rigid registration. In the present implementation, the images, including the adaptive block size, were used on the same scale. If the same scale was maintained throughout the three levels of block sizes, the larger block may help prevent spurious large displacements because it has more data to guide the algorithm. The downside is that a large block may do a poor job of capturing distortions that are not uniform over the block. To reduce this problem, one popular approach is to used multi-scale images (e.g. Gaussian pyramid), each level being twice as coarse in resolution as the previous, which has been shown effective in reducing computation time [225,232]. However, no comparison was made to study the difference of registration accuracy with and without subsampling after Gaussian filtering in [225,232]. In the present study little, if any, difference was found between default implementation with and without filtered subsampling. This may due to the fact that the boxcar filter is not as effective as a Gaussian filter to avoid local maxima or the effect of image blurring to a lower scale, together with MSCR, has little effect on

registration accuracy. Further experiments are needed to study the effectiveness of filtered subsampling strategy and any possible effect of scale values of a Gaussian kernel on MSCR. The scale parameters may tie to the surrounding voxels so that they vary spatially [100].

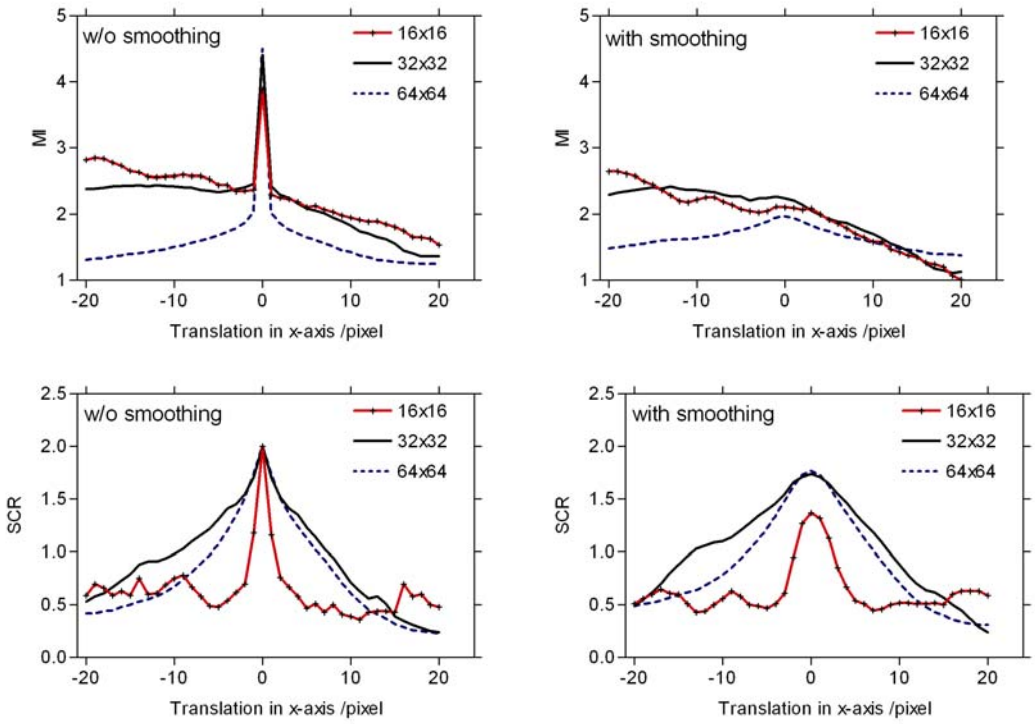


Figure 8.14: MI (top) and SCR (bottom) as a function of translation along the x-axis, for the registration of the CT-CT image blocks of size: 16×16, 32×32 and 64×64. Data with and without (w/o) smoothing (accomplished by filtering one of the CT images with a boxcar filter) are plotted separately. The zero position corresponds to the correct registration. In all cases, full sampling and 256 intensity bins were used.

A qualitative examination of the registration in the brain region demonstrates that MSCR is robust and accurate for high resolution image volume pairs. For example, 91% of image voxels were registered within an accuracy of 2 voxels in both T_1-T_1 and PD-T_2 MR image registration, which represents 42% and 30% displacement error improvement, respectively (Figure 8.6 and 8.11). However, a significant difference in resolution and intensity distribution between the two images (e.g. CT-SPET, MR-SPET), or images of low resolution, profoundly deteriorates the registration accuracy and robustness. The improvement for SPET-MR registration is marginal (about 10%). The results suggest that common anatomical information, high spatial resolution and low noise level may have a significant effect on the performance of registration. The poor performance of the non-rigid algorithm in brain SPET-MR registration and torso SPET-CT registration may be attributed to poor resolution and counting statistics of SPET

compared to CT and MR. Examples of the MI and SCR are illustrated in Figure 8.14 when a CT image block was translating with respect to another CT image block along x-axis. The image block, showed in Figure 8.15 (top row), is registered as a function of translation to a second image block derived from the identical CT with and without blurring. With smoothing, a 5×5 pixels boxcar-smoothing filter was applied to the CT image to mimic resolution difference between images for registration.

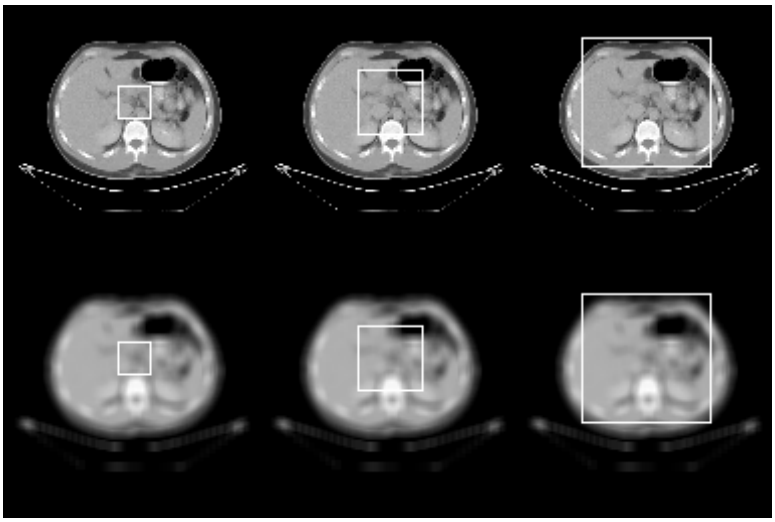


Figure 8.15: Image blocks (enclosed by windows), with (bottom) and without (top) smoothing, correspond to sizes (from left to right): 16×16, 32×32 and 64×64. These sub-images are used to compute MI and SCR as a function of translation along x-axis for two cases: CT-CT and CT-blurred-version-CT.

The results in Figure 8.14 show that MI, NMI (not shown here) and SCR gave correct registration at zero position when high resolution CT image blocks were registered regardless the size of the block. This further supports, with MSCR algorithm, that overlapping image blocks are able to give optimum local alignment at correct registered positions for high resolution images registration using local translation. When reducing the sizes of the image blocks progressively from 64^2 to 16^2, MI appeared not symmetrical about the zero translation, and 16^2 suffered the most. Similar observations were found for NMI. Based on these results, the size of image block had no effect on SCR. When a smoothing filter was applied to one of the CTs prior to registration, peaks of MI and SCR at registered positions, to some extent, were suppressed and broadened with respect to the corresponding maxima without blurring. Surprisingly, with block sizes at 32^2 and 16^2, the optimal registration at zero position disappeared for MI but not for SCR. This demonstrates that, with image blurring, reducing image block size affects MI and NMI (not shown) by concealing the function peaks and, therefore, deteriorates local registration. However, the sensitivity and severity of this effect is not uniform

across similarity measures. MI and NMI did suffer most while SCR was unaffected although local maxima are observed at the smallest block size 16^2 for SCR. This observation helps to explain the poor results obtained for local registration between image pairs with a large difference in resolution (e.g. SPET-MR and SPET-CT) comparing to high resolution image pairs (e.g. MR-MR) local alignment.

Furthermore, reducing the sub-image volume for progressively overlapping smaller image blocks also implies less image information content for registration. The low resolution, difference in the context of information (i.e. functional versus structural), absence of anatomical detail and paucity of compatible corresponding structures between SPET and CT/MR images compounds deterioration towards similarity measures on maxima, resulting in misregistration. Thus similarity measures derived from these image blocks tend to be poor and there is no "true" global optimum to drive the match and so the registration is prone to be trapped at a "false" local optimum. Another contributory factor in the poor registration may be that sharp edges at some boundaries produce a profound effect on the joint intensity histogram for registration. For example, sharp edges on the surface of the liver in SPET may overshoot the corresponding surface in the CT image and instead align with the body surface (Figure 8.13). Such results indicate that the edges of more global structures in SPET (e.g. image boundary) may provide sufficiently strong influence on the joint histogram to drive the non-rigid registration. To verify this, a CT sub-image is translated along the x-axis with respect to an identical CT sub-image blurred with a boxcar filter, and MI was computed as a function of translation (Figure 8.16).

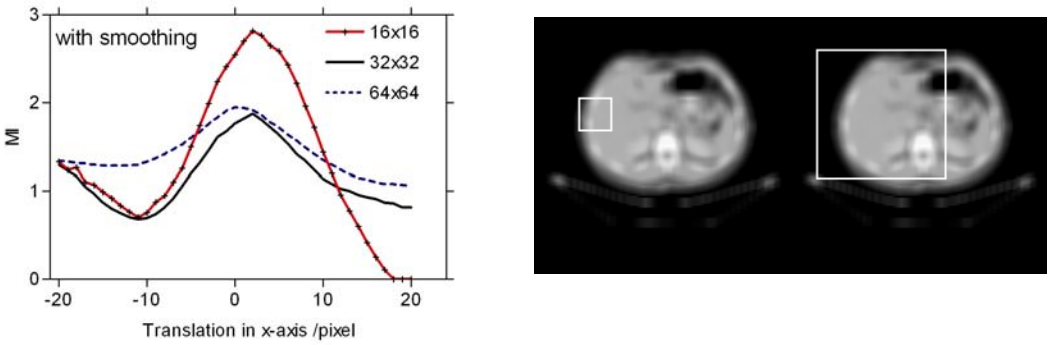


Figure 8.16: Variation of MI as a function of translation of a CT image block to an identical CT blurred with a 5^2 pixels boxcar filter (left). The zero translation corresponds to the registration position. The windows in the CTs enclose sub-images of sizes at 16^2 (middle) and 64^2 (right) pixels. In both cases, image boundary is deliberately included for the registration.

The image boundary is deliberately included as a global anatomical structure in the sub-image for registration using full sampling and 256 intensity bins. For SCR, correct registration at zero translation was obtained for all three sub-image blocks. In contrast to Figure 8.14 (top right), optimal peaks are clearly apparent irrespective of the image block sizes although the peaks of 16^2 and 32^2 are slightly shifted to the positive x-direction. This signifies that the body boundary plays an important role to constrain image registration and drives global alignment in image block registration.

The visual assessment of colour overlay display for SPET-CT/MR registration is intuitive and direct; it is sensitive enough to detect a small mismatch along edges of anatomical structures. However, it is difficult to quantify the degree of mismatch. The unsatisfactory results for SPET-CT registration in the thoracic-abdominal region may suggest a shortfall of the non-rigid registration in low-resolution images where only limited information is available within most image blocks. Therefore, use of the full 3D image content is a possible alternative approach for non-rigid registration on low-resolution images. Based on the mutual information as a global similarity measure, Kim *et al.* [380] demonstrated an automated thin-plate spline warping to register FDG images of a rat brain with its video reference images. In this method, an arbitrary set of landmarks is supplied initially and then the landmarks are iteratively repositioned until the resulting warp optimises the global mutual information. Meyer *et al.* [186,381] presented a thin-plate spline warped registration of thoracic PET-CT and volumetric ultrasound scans where control points changed iteratively to maximize mutual information. In these studies, the full image volumetric data were used to compute a global similarity measure to register low-resolution image non-rigidly.

The MSCR algorithm is based on the warping described by the displacement of regular grid nodes. It relies on the sampling of a displacement field at a set of nodes, each node anchored at the center of an image block. Gaussian interpolation is then used to give displacements at inter-nodal locations. In order to recover the distortion exactly, the grid spacing (i.e. reciprocal of spatial sampling rate) must be no greater than one-half the period of the finest variation (corresponding to the maximum spatial frequency) within the displacement field [262]. In this case, the displacement field is said to be sampled at its Nyquist rate, according to the sampling theorem. If the inter-nodal spacing is greater than required by the Nyquist criterion, aliasing will result. Further

work is needed to assess the influence of the rate at which the displacement vector field is sampled on the displacement error. Although sampling theorem provides for maximum sampling interval but not minimum, it may be argued that each sample represents a finite portion of the displacement field. Thus smaller grid spacing may be necessary to capture mismatches at a finer level. Again this is a good subject for further work.

The symmetric correlation ratio gives comparable registration to mutual information in 3D intra-modality and inter-modality tasks at full sampling but it is superior to MI in registering sparsely sampled high resolution MR images. The sparse sampling refers to the coarse to fine strategy used to sample the displacement field. SCR performed virtually identically to normalised mutual information (NMI). For the non-rigid registration of SPET-MR in the brain and SPET-CT in torso, none of the similarity measures (SCR, MI and NMI) led to an effective registration. One possible reason for the failure may be due to the very different intensity distributions in SPET and CT/MR. In fact, the image block approach together with very poor resolution of one of the images often reduces information content. This substantially increases the dispersion of joint intensity histogram. One way to circumvent this problem is to register the transmission (TR) image with the CT image and apply the optimum transformation to the SPET image. In this case, both TR and CT images represent an attenuation map and therefore have comparable intensity distributions although TR has an inferior spatial resolution [382]. Intensity transformation between two images prior to registration is also suggested in [370, 372] to improve the performance of inter-modality registration.

The MSCR algorithm performed poorly with respect to all studies with the low-resolution image because of the physical limitation and constraints in clinical imaging conditions. Non-rigid registration techniques developed for the high-resolution image pair does not extend effectively to the characteristically low-quality and noisy SPET images of deformable organs. In the case of a pathological reduction or absence of radiotracer uptake, the correspondence between organs visualized in functional and anatomical images may be impaired. As body organs are not rigid and are closely packed in the thoracic or abdominal space, a phantom or a simulated local distortion may not be representative of the complex distortion in the abdominal-thoracic region in clinical setting. Physical or numerical phantoms studies provide insight in the

performance of the algorithm but little is known in terms of the performance for real clinical data. The models also do not allow the exploration of the algorithms under a variety of conditions anticipated in diagnostic and therapeutic applications, nor is it possible to discriminate between intensity artefacts, geometric distortion, patient-specific abnormalities, and normal variations, such as different liver positions (high/low) and bladder content (from full to empty).

8.4 Conclusion

This chapter assesses the feasibility of using a symmetric correlation ratio and sub-image block matching technique to spatially register 3D intra- and inter-modality images. For the high-resolution image, rich in information content, the MSCR algorithm performs satisfactorily. The accuracy of the technique was evaluated on digital phantoms and on patient data. In both studies, registration was successful for 92% of image voxels with an accuracy of ≤ 2 voxels. A high success rate shows that the MSCR method is reliable and robust. With the default implementation, the mean displacement error was less than the MR voxel for intra-modality MR registration. In the current implementation, the registration process takes about 10 minutes to complete a 3D alignment for image pairs of $128 \times 128 \times 99$ voxels on a stand alone Pentium IV PC with 1.7 GHz CPU and 256 Mbytes random access memory on board. Further investigation is warranted to examine aspects of non-rigid registration between low resolution image (e.g. SPET) and high resolution image (e.g. CT, MR).

Chapter 9

Conclusions and further work

The ability to register medical images from different modalities is critically important in clinical applications [47,49,50,51,52,56], such as radiological diagnosis [53,54], treatment planning prior to surgery [60,67,68,69,71] and radiotherapy [70,72,73,74], and monitoring response to treatment [58,65,66]. Differences in positioning, image quality and scanning protocol, as well as patient motion combine to make registration nontrivial. The registration problem is especially difficult if one image represents functional information while the other represents anatomical information. This challenge has spurred the development of many registration algorithms in the last two decades [383]. In mid 1990s, new intensity-based similarity measures based on entropy [176,178,180] and local variance [172,131,181] received much attention especially in multimodality image registration. These intensity-based measures proved to be quite robust across a range of 3D image registration problems. They have been successfully applied to a variety of problems and have been shown, using a brain dataset with known ground truth, to be superior compared to feature-based algorithms [122,123]. In this thesis, multi-level rigid and non-rigid registration is introduced and discussed. Their implementation, validation and clinical applications are also demonstrated for rigid-body registration. Similarity measures based on the joint intensity histogram (MI, NMI and SCR) have been used in fully automated image registration algorithms without the need for pre-registration processing. At the outset of this project, many aspects of these measures were not fully understood. How do they differ in performance? What is the best way of implementing these measures? What is the best strategy to accelerate the registration process? Is pre-processing of images beneficial? For what type of image are they successful and for which images do they fail? Can the registration process be extended to non-rigid registration? How can these measures be applied to register images outside the brain? In this thesis, intensity-based strategies have been examined in detail and refinements introduced to further improve registration performance. Emphasis has been given to methods of evaluating and validating these registration strategies with particular attention paid to the registration of low resolution emission tomographic data. The development of realistic simulations based on incorporating clinical data in a thoracic-abdominal phantom was central to the validation. The

techniques were further incorporated in the development of non-rigid registration algorithms.

9.1 Rigid body registration

In the rigid-body registration, the question of whether the choice of the similarity measure has a bearing on registration accuracy was investigated with clinical data (Chapter 4) and with simulation of a thoracic-abdominal SPET (Chapter 6). The comparison was carried out over a wide range of image conditions, acquisition regimes, reconstruction and registration strategies. Three intensity-based similarity measures (SCR, NMI and MI) were compared in rigid-body SPET-MR/CT registration using a multi-level strategy. In clinical images, external fiducial markers were used as a reference for the evaluation of registration accuracy. For simulated SPET, registration accuracy was computed from the discrepancy between simulated and applied transformations averaged over the eight corners of a cuboid bounding box. In some cases, the mean registration error and its standard deviation were strongly biased by outliers as a result of the registration algorithm terminating at a spurious local optimum. The size of the error for an outlier is not very interesting from the point of view of registration performance; the registration can be considered a failure whether the error is, say, 150 mm or 175 mm. To assess the performance of registration from residual errors, a registration with an error exceeding 100 mm was considered a failure and was not included in the computation of the summary statistics. This value was set empirically because errors greater than 100 mm were found in some cases. A measure of the robustness of the algorithm was provided by the success rate (proportion of registrations with the mean error below a set threshold). A difficulty with this measure is the need to set the threshold, which ultimately should be related to the accuracy required by each clinical application. In the present studies, 20 mm was used as the threshold to compute the success rate.

Overall, the results demonstrate that the performance of registration depends on the type of SPET image and implementation strategy. The results were not the same for the three similarity measures; MI was more susceptible to outliers than SCR and NMI when sparse sampling was used for registration. In registering SPET ^{99m}Tc-HMPAO studies to MR at full sampling (Chapter 4), the registration performance of all three similarity measures was statistically indistinguishable. Registration at full sampling would take

too long for clinical use. Acceleration was accomplished with a multi-level sampling strategy and adaptive intensity rebinning. Substantial acceleration was achieved without adversely affecting accuracy. With SCR and NMI, but not MI, introducing multi-level sampling did not significantly change the registration accuracy. In terms of the mean registration error measured relative to the fiducial marks, SCR performed better (3.9±1.1 mm) than MI (5.6±2.3 mm), as confirmed by statistical testing at the 5% confidence level. However, the mean errors for MI (5.6±2.3 mm) and NMI (4.4±1.1 mm) were not significantly different. In the present implementation, an acceleration factor of about 20 was achieved for 3D images. The mean error for SCR and the multi-level sampling (3.9±1.1 mm) is comparable to results of Collignon (3.6 mm) and Hill (3.2 mm) reported for PET-MR registration [122].

With simulated thoracic-abdominal SPET (Chapter 6), SCR also performed better than MI and NMI in the registration of MDP FBP-reconstructed SPET and CT. However, registration accuracy was found to be similar for simulated ^{67}Ga SPET-CT registration using NMI, MI and SCR irrespective of which reconstruction algorithm (OSEM and FBP) was used, with and without attenuation compensation. Since image quality is different for FBP and OSEM, the assessment of registration performance was based on noise-equivalent reconstructions. When noise-equivalent reconstructions were selected, registration accuracy did, to some extent, depend on the choice of the reconstruction algorithm as well as on the use of AC. Registration of FBP-reconstructed SPET was more likely to be affected by AC than the OSEM-reconstructed SPET. With AC, OSEM-reconstructed SPET gave better performance than a noise-equivalent FBP reconstruction. Without AC, the effect was less apparent. Registration deteriorated for ^{67}Ga SPET reconstructed by FBP with AC (Subsection 6.4.1). However, no significant difference was found for OSEM-reconstructed SPET with and without AC. A proposition that registration accuracy depends on the reconstruction noise was not supported by the experiments. Within reasonable limits, changing the cutoff frequency in FBP and the number of iterations in OSEM had little, if any, effect on registration. This signifies that variance and entropy based algorithms are insensitive to image reconstruction noise. It was postulated that a number of practical factors, namely matching CT resolution to that of SPET, introducing a distinct body boundary in SPET, matching the field of view between modalities and using a large image volume could

improve registration performance but extensive experimentation failed to demonstrate such effects. One exception was the registration of OSEM-reconstructed ^{67}Ga SPET, where the resolution adjustment was associated with a substantial increase in the number of failures (mean error ≥ 20 mm). For OSEM-reconstructed ^{67}Ga SPET, the failure increased from 0 to 12% (SCR), 0 to 32% (NMI) and 4% to 28% (MI), when resolution matching was applied to CT. Experiments also suggested that CT intensity windowing can improve the performance of registration for OSEM ^{67}Ga SPET but little effect was found for MDP SPET. This means that its effect on registration depends on the SPET image. Furthermore, results suggest that using a very small image volume for registration reduces performance. Registration accuracy of about 11.0 mm for thoracic-abdominal region was obtained using SCR. It is worth to note that this error is similar to the result for PET-CT registration (10.6 mm) reported by Skalski *et al.* [351]. This size of error was comparable to the registration error obtained when the CT image was registered to a low-resolution version of itself (about 11.0 mm) but was much larger than the error (about 2.0 mm) arising when the CT image was registered to itself at its original resolution. This indicates that one of the major factors limiting the registration accuracy may be image resolution (Subsection 6.5.6).

9.2 Non-rigid body registration

A non-rigid image registration method has been examined for inter-subject intra-modality brain image (Subsection 8.2.1-8.2.2) and intra-subject inter-modality thoracic-abdominal regions (Subsection 8.2.3). The method constructs a displacement vector field by finding a translation vector that maximizes SCR (or NMI, MI) for each of a set of overlapping image blocks. The method employs a multi-level strategy to provide consecutive refinements in block size. A median filter is used to remove outliers in the displacement field and a Gaussian function smoothly interpolates the measured displacement field. The question of whether the choice of the similarity measure affects the registration performance was investigated using 3D clinical data for MR-MR and MR-SPET registration of the brain region, and simulated SPET for CT-SPET registration of thoracic-abdominal region. Known non-rigid misalignments were applied to the clinical images and the simulated images. For intra-modality and "inter-modality" registration of MR brain images, SCR gave 92% mean success rate over ten independent measurements, where the success threshold was set to the error of 2 voxels. The results also demonstrated that SCR gave a comparable performance to MI and NMI,

but, as in the case of rigid registration, SCR performed better for sparsely sampled data. However, in the MR-SPET and CT-SPET non-rigid registration of the non-brain region, none of the three measures was able to recover local distortions.

9.3 Further work

Although intermodality registration of images has been largely successful in the brain, the thoracic-abdominal region presents a greater challenge because changes in patient pose between procedures can induce greater variation in the volume, shape and position of constituent organs. To evaluate rigid registration, image pairs were arbitrarily misaligned by 6 parameters. These transformations might not be clinically realistic. Further work is needed to validate the algorithm with real distortions. Indeed, only clinical studies can determine usefulness of a registration and required registration accuracy for specific clinical applications. In order to ensure the relevance of the registration results to a particular clinical situation, prospective studies must first be performed to determine the nature and extent of distortion appearing in each clinical procedure. These distortions should represent a typical range of misalignment encountered in clinical applications. The residual RMS distance between corresponding points of interest (i.e. the target registration error, TRE) appears more relevant to clinical application than distances derived from the corner points of the bounding box. TRE depends very much on the ability to accurately localize abnormal activity uptake in SPET to an anatomical structure in CT/MR. Ambiguities concerning final interpretation may be resolved by a visual assessment of fused SPET and CT/MR images. Thus, further studies should include visual verification by clinical specialists to provide additional assessment of the registration performance. Local misalignment arises not only from changes in patient pose and movement during acquisition but also from the involuntary motion of internal organs. In addition, local misalignment can occur between image acquisitions due to pathological changes or therapeutic intervention (e.g. surgery). In intersubject applications, morphological differences between subjects can also be represented as local misalignments. The multi-level algorithm was successful in high resolution MR-MR non-rigid registration but not in SPET-MR and SPET-CT registrations. This failure may be due to image block becoming too small. Moreover, failure may arise from the differences in implementation between MR-MR and SPET-CT/MR registration. Extension to intra-subject non-rigid registration between high resolution images, such as CT or MR images, and low resolution images, such as

SPECT or PET images, or low resolution to low resolution images, requires further investigation. Improvements in the multi-level non-rigid algorithm may involve accommodating block size and other implementation parameters to image resolution. The inter-modality registration may be improved further by the use of *a priori* training models [208,384] where prior knowledge of the joint intensity histogram from a similar pair of registered images is used to constrain registration. Another approach is to register the transmission image (TR), acquired immediately before or after the SPET scan, with the CT image, and apply the resulting displacement vector field to the SPET image. The TR scan and the emission image are assumed to be registered. The TR and the CT image are both based on the same physical attenuation process but differ in the energy spectrum of the radiation and the type of detector used [382]. The TR-CT registration can be regarded as an intra-modality registration problem, which could be less difficult than the inter-modality registration. Non-rigid registration is difficult to validate without a gold standard. Therefore, one possible approach to validation is to use an image pair from a dual-modality system (e.g. dedicated PET/CT scanner [385,386,387]) as registered images. Clinically realistic misalignment could then applied to one of the images and the other image would be registered to the misaligned image. The registration accuracy could be computed from the difference between the recovered and the applied misalignment. It may be supposed that the advent of dual-modality systems obviates the need for software based solutions to registration. However, such systems are costly and still uncommon. Moreover, software solutions are still required to account for a range of misalignments (e.g. those arising from the differences in acquisition time in SPET/PET and CT/MR) [76,388] Incidentally, the mere fact that manufacturers introduced dual modality systems is testament to the recognition of the need for clinical registration.

Investigation is also required to optimize further the performance of the proposed algorithm. For example, a strict scale space representation of images may improve convergence properties. Furthermore, the focus should be on additional investigation of the application of intensity-based measures to non-brain multimodality registration. The research in non-rigid registration still has many difficult challenges ahead, both in terms of addressing the practical needs of clinical applications as well as developing appropriate validation and evaluation approaches across different image modalities. A key challenge here is to understand what possible local deformations are feasible for

various organs, such as the lungs and the spine. This knowledge could form the basis of a deformation model for particular organs. Such models may serve to constrain the set of possible transformations. Many of these problems require the analysis of not only the normal structure and function, but also of different abnormal states. Further attention needs to be given to the anatomical structures in functional images that critically influence registration accuracy at a global or local level. For example, the spine and the kidneys are well visualized on some functional scans, but the external body outline and the lungs are not, although they are well visualized on the anatomical scans. Where the information present in the functional image is insufficient to support registration with an anatomical image, additional information is required. Some research groups have already made progress in this direction by incorporating a second radionuclide [81], transmission scans [260,389,390] or scatter window data [254] in image registration. More effort is needed to investigate how this extra information can be combined globally or locally across different data sets for better alignment.

Bibliography

[1] A. L. Weber, History of Head and Neck Radiology: Past, Present, and Future, *Radiology* **218**, 15-24, 2001.

[2] P. N. T. Wells, Physics and engineering: milestones in medicine, *Medical Engineering & Physics* **23**, 147-153, 2001.

[3] G. N. Hounsfield, Computerized transverse axial scanning (tomography), Part I: description of system, *Brit J Radiol* **46**, 1016-1023, 1973.

[4] Y. D'Asseler, M. Koole, K. Van Laere, S. Vandenberghe, L. Bouwens, R. Van de Walle, C. Van de Wiele, I. Lemahieu, and R. A. Dierckx, PACS and multimodality in medical imaging, *Technol Health Care* **8**, 35-52, 2000.

[5] H. Pohjonen, Image fusion in open-architecture PACS-environment, *Comput Methods Programs Biomed* **66**, 69-74, 2001.

[6] R. A. Brooks and G. Di Chiro, Principles of computer-assisted tomography in radiographic and radioisotopic imaging, *Phys Med Biol* **21**, 689-732, 1976.

[7] S. Webb, *The physics of medical imaging*, Bristol: Adam Hilger, 1988.

[8] M. M. Ter-Pogossian, Basic principles of computed axial tomography, *Semin Nucl Med* **7**, 109-127, 1977.

[9] J. E. Barnes, Characteristics and control of contrast in CT, *Radiographics* **12**, 825-837, 1992.

[10] U. Meyer, T. Stamm, N. Meier, and U. Joos, First experience with a public domain computer-aided surgical system, *Br J Oral Maxillofac Surg* **40**, 96-104, 2002.

[11] Y. Hagiwara, M. Koizumi, and T. Igarashi, Application of CT imaging for dental implant simulation, *J Oral Sci* **41**, 157-161, 1999.

[12] H. Handels, J. Ehrhardt, W. Plotz, and S. J. Poppl, Simulation of hip operations and design of custom-made endoprostheses using virtual reality techniques, *Methods Inf Med* **40**, 74-77, 2001.

[13] Y. Okajima, H. Tajima, T. Kumazaki, and M. Onda, Clinical application of a CT-guided lung biopsy system: core needle biopsy at the IVR center, *J Nippon Med Sch* **69**, 434-444, 2002.

[14] F. J. Ruhli, J. Hodler, and T. Boni, Technical note: CT-guided biopsy: A new diagnostic method for paleopathological research, *Am J Phys Anthropol* **117**, 272-275, 2002.

[15] N. Archip, P. J. Erard, J. M. Haefliger, and J. F. Germond, A computer aided diagnostic system for radiotherapy planning, *Med Phys* **12**, 246-251, 2002.

[16] T. Ishigaki, S. Sakuma, T. Banno, M. Imazawa, T. Tanaka, and K. Araki, Computer-assisted conformation radiotherapy system. An application of CT, *Eur J Radiol* **3**, 367-374, 1983.

[17] W. A. Edelstein, J. M. S. Hutchison, G. Johnson, and T. Redpath, Spin-warp NMR imaging and applications to human whole-body imaging, *Phys Med Biol* **25**, 751-756, 1980.

[18] D. R. Bailes and D. J. Bryant, NMR imaging, *Contemp Phys* **25**, 441-475, 1984.

[19] G. A. Wright, Magnetic resonance imaging, *IEEE Signal Processing Magazine* **14**, 56-66, 1997.

[20] R. K.-S. Kwan, A. C. Evans, and G. B. Pike, MRI simulation-based evaluation of image-processing and classification methods, *IEEE Trans Med Imag* **18**, 1085-1097, 1999.

[21] F. Bloch, Nuclear induction, *Phys Rev* **70**, 460-474, 1946.

[22] E. R. Andrew, Nuclear magnetic resonance and the brain, *Brain Topogr* **5**, 129-133, 1992.

[23] E. M. Haacke and J. A. Tkach, Fast MR imaging: Techniques and Clinical Applications, *Am J Roentgenology* **155**, 951-964, 1990.

[24] P. C. Lauterbur, Imaging formation by induced local interactions: examples employing nuclear magnetic resonance, *Nature* **242**, 190-191, 1973.

[25] M. Filippi, Linking structural, metabolic and functional changes in multiple sclerosis, *Eur J Neurol* **8**, 291-297, 2001.

[26] S. J. Hickman and D. H. Miller, Imaging of the spine in multiple sclerosis, *Neuroimaging Clin N Am* **10**, 689-704 ,viii, 2000.

[27] Y. X. Wang, S. M. Hussain, and G. P. Krestin, Superparamagnetic iron oxide contrast agents: physicochemical characteristics and applications in MR imaging, *Eur Radiol* **11**, 2319-2331, 2001.

[28] V. M. Runge, A review of contrast media research in 1999-2000, *Invest Radiol* **36**, 123-130, 2001.

[29] M. Regenfus, D. Ropers, S. Achenbach, W. Kessler, G. Laub, W. G. Daniel, and W. Moshage, Noninvasive detection of coronary artery stenosis using contrast-enhanced three-dimensional breath-hold magnetic resonance coronary angiography, *J Am Coll Cardiol* **36**, 44-50, 2000.

[30] D. C. Noll, C. R. Genovese, L. E. Nystrom, A. L. Vazquez, S. D. Forman, W. F. Eddy, and J. D. Cohen, Estimating test-retest reliability in functional MR imaging. II: Application to motor and cognitive activation studies, *Magn Reson Med* **38**, 508-517, 1997.

[31] H. O. Anger, Scintillation camera, *Rev Sci Instrum* **29**, 27-33, 1958.

[32] T. F. Budinger and G. T. Gullberg, Three-dimensional reconstruction in nuclear medicine emission imaging, *IEEE Trans Nucl Sci* **21**, 21-80, 1974.

[33] O. Bonne, Y. Krausz, and B. Lerer, SPECT imaging in psychiatry. A review, *Gen Hosp Psychiatry* **14**, 296-306, 1992.

[34] J. C. Masdeu, L. M. Brass, B. L. Holman, and M. J. Kushner, Brain single-photon emission computed tomography, *Neurology* **44**, 1970-1977, 1994.

[35] M. A. King, B. M. W. Tsui, and T.-S. Pan, Attenuation compensation for cardiac single-photon emission computed tomographic imaging: part 1. Impact of attenuation and methods of estimating attenuation maps, *J Nucl Cardiol* **2**, 513-524, 1995.

[36] M. A. King, B. M. W. Tsui, T.-S. Pan, S. J. Glick, and E. J. Soares, Attenuation compensation for cardiac single-photon emission computed tomographic imaging: part 2. Attenuation compensation algorithms, *J Nucl Cardiol* **3**, 55-63, 1996.

[37] B. M. W. Tsui, X. D. Zhao, G. K. Gregoriou, D. S. Lalushl, E. C. Frey, R. E. Johnston, and W. H. McCartney, Quantitative cardiac SPECT reconstruction with reduced image degradation due to patient anatomy, *IEEE Trans Nucl Sci* **41**, 2838-2844, 1994.

[38] B. F. Hutton, H. M. Hudson, and F. J. Beekman, A clinical perspective of accelerated statistical reconstruction, *Eur J Nucl Med* **24**, 797-808, 1997.

[39] R. E. Coleman, J. M. Hoffman, M. W. Hanson, H. D. Sostman, and S. C. Schold, Clinical application of PET for the evaluation of brain tumors, *J Nucl Med* **32**, 616-622, 1991.

[40] P. S. Conti, D. L. Lilien, K. Hawley, J. Keppler, S. T. Grafton, and J. R. Bading, PET and [18F]-FDG in oncology: a clinical update, *Nucl Med Biol* **23**, 717-735, 1996.

[41] M. M. Ter-Pogossian, N. A. Mullani, D. C. Ficke, J. Markham, and D. L. Snyder, Photon time-of-flight-assisted positron emission tomography, *J Comput Assist Tomogr* **5**, 227-239, 1981.

[42] G. Muehllehner and J. S. Karp, Positron emission tomography imaging-technical considerations, *Sem Nucl Med* **16**, 35-50, 1986.

[43] N. D. Volkow, N. A. Mullani, and B. Bendriem, Positron emission tomography instrumentation: an overview, *Am J Physiol Imag* **3**, 142-153, 1986.

[44] K. Ishii, Clinical application of positron emission tomography for diagnosis of dementia, *Ann Nucl Med* **16**, 515-525, 2002.

[45] C. Messa, F. Fazio, D. C. Costa, and P. J. Ell, Clinical brain radionuclide imaging studies, *Semin Nucl Med* **25**, 111-143, 1995.

[46] D. J. Hawkes and D. L. G. Hill, Medical imaging at Guy's Hospital, King's College London, *IEEE Trans Med Imag* **22**, 1033-1041, 2003.

[47] J. B. Maintz and M. A. Viergever, A survey of medical image registration, *Med Image Anal* **2**, 1-36, 1998.

[48] D. L. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes, Medical image registration, *Phys Med Biol* **46**, R1-45, 2001.

[49] B. F. Hutton, M. Braun, L. Thurfjell, and D. Y. Lau, Image registration: an essential tool for nuclear medicine, *Eur J Nucl Med Mol Imaging* **29**, 559-577, 2002.

[50] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, Mutual-information-based registration of medical images: a survey, *IEEE Trans Med Imag* **22**, 986-1004, 2003.

[51] G. T. Chen, C. A. Pelizzari, and D. N. Levin, Image correlation in oncology, *Important Adv Oncol*, 131-141, 1990.

[52] D. A. Weber and M. Ivanovic, Correlative image registration, *Semin Nucl Med* **24**, 311-323, 1994.

[53] P. G. Kluetz, C. C. Meltzer, V. L. Villemagne, P. E. Kinahan, S. Chander, M. A. Martinelli, and D. W. Townsend, Combined PET/CT imaging in Oncology: Impact on patient management, *Clinical Positron Imaging* **3**, 223-230, 2000.

[54] R. B. Howman-Giles, D. K. Chung, Y. H. Lau, T. Kitsos, R. Uren, and B. F. Hutton, Does mulit-modality image co-registration impact on clinical management in paediatric diseases?, *J Nucl Med* **43(5S)**, 84-85, 2002.

[55] B. Fei, C. Kemper, and D. L. Wilson, A comparative study of warping and rigid body registration for the prostate and pelvic MR volumes, *Computerized Medical Imaging and Graphics* **27**, 267-281, 2003.

[56] P. A. van den Elsen, E.-J. D. Pol, and M. A. Viergever, Medical image matching: a review with classification, *IEEE Eng Med Biol* **12**, 26-39, 1993.

[57] T. J. O'Brien, M. K. O'Connor, B. P. Mullan, B. H. Brinkmann, D. Hanson, C. R. Jack, and E. L. So, Subtraction ictal SPET co-registered to MRI in partial epilepsy: description and technical validation of the method with phantom and patient studies, *Nucl Med Commun* **19**, 31-45, 1998.

[58] P. J. Lewis, A. Siegel, A. M. Siegel, C. Studholme, J. Sojkova, D. W. Roberts, V. M. Thadani, K. L. Gilbert, T. M. Darcey, and P. D. Williamson, Does performing image registration and subtraction in ictal brain SPECT help localize neocortical seizures?, *J Nucl Med* **41**, 1619-1626, 2000.

[59] C. Perault, D. Papathanassiou, H. Wampach, P. Vera, A. Kaminska, C. Chiron, P. Peruzzi, and J. C. Liehn, Computer-aided intrapatient comparison of brain SPECT images: the gray- level normalization issue applied to children with epilepsy, *J Nucl Med* **43**, 715-724, 2002.

[60] M. Murphy, T. J. O'Brien, K. Morris, and M. J. Cook, Multimodality image-guided epilepsy surgery, *J Clin Neurosci* **8**, 534-538, 2001.

[61] A. R. Giacometti, P. C. Davis, N. P. Alazraki, and J. A. Malko, Anatomic and physiologic imaging of Alzheimer's disease, *Clin Geriatr Med* **10**, 277-298, 1994.

[62] C. DeCarli, The role of neuroimaging in dementia, *Clin Geriatr Med* **17**, 255-279, 2001.

[63] K. Ishii, F. Willoch, S. Minoshima, A. Drzezga, E. P. Ficaro, D. J. Cross, D. E. Kuhl, and M. Schwaiger, Statistical brain mapping of 18F-FDG PET in Alzheimer's disease: validation of anatomic standardization for atrophied brains, *J Nucl Med* **42**, 548-557, 2001.

[64] P. J. Slomka, P. Radau, G. A. Hurwitz, and D. Dey, Automated three-dimensional quantification of myocardial perfusion and brain SPECT, *Comput Med Imaging Graph* **25**, 153-164, 2001.

[65] S. J. Nelson, M. R. Day, P. J. Buffone, L. L. Wald, T. F. Budinger, R. Hawkins, W. P. Dillon, S. Huhn, M. D. Prados, S. Chang, and D. B. Vigneron, Alignment of volume MR images and high resolution [18F]fluorodeoxyglucose PET images for the evaluation of patients with brain tumors, *J Comput Assist Tomogr* **21**, 183-191, 1997.

[66] R. Myers, The application of PET-MR image registration in the brain, *Br J Radiol* **75 Spec No**, S31-5, 2002.

[67] K. R. Smith, K. J. Frank, and R. D. Bucholz, The NeuroStation--a highly accurate, minimally invasive solution to frameless stereotactic neurosurgery, *Comput Med Imaging Graph* **18**, 247-256, 1994.

[68] D. L. Hill, D. J. Hawkes, M. J. Gleeson, T. C. Cox, A. J. Strong, W. L. Wong, C. F. Ruff, N. D. Kitchen, D. G. Thomas, A. Sofat, and et al., Accurate frameless registration of MR and CT images of the head: applications in planning surgery and radiation therapy, *Radiology* **191**, 447-454, 1994.

[69] C. A. Pelizzari, Image processing in stereotactic planning: volume visualization and image registration, *Med Dosim* **23**, 137-145, 1998.

[70] J. G. Rosenman, E. P. Miller, G. Tracton, and T. J. Cullip, Image registration: an essential part of radiation therapy treatment planning, *Int J Radiat Oncol Biol Phys* **40**, 197-205, 1998.

[71] J. Kettenbach, T. Wong, D. Kacher, N. Hata, R. B. Schwartz, P. M. Black, R. Kikinis, and F. A. Jolesz, Computer-based imaging and interventional MRI: applications for neurosurgery, *Comput Med Imaging Graph* **23**, 245-258, 1999.

[72] C. C. Teng, M. M. Austin-Seymour, J. Barker, I. J. Kalet, L. G. Shapiro, and M. Whipple, Head and Neck Lymph Node Region Delineation with 3-D CT Image Registration, *Proc AMIA Symp*, 767-771, 2002.

[73] S. Mutic, J. F. Dempsey, W. R. Bosch, D. A. Low, R. E. Drzymala, K. S. Chao, S. M. Goddu, P. D. Cutler, and J. A. Purdy, Multimodality image registration quality assurance for conformal three- dimensional treatment planning, *Int J Radiat Oncol Biol Phys* **51**, 255-260, 2001.

[74] M. T. Munley, L. B. Marks, C. Scarfone, G. S. Sibley, E. F. Patz, Jr., T. G. Turkington, R. J. Jaszczak, D. R. Gilland, M. S. Anscher, and R. E. Coleman, Multimodality nuclear medicine imaging in three-dimensional radiation treatment planning for lung cancer: challenges and prospects, *Lung Cancer* **23**, 105-114, 1999.

[75] E. J. Somer, P. K. Marsden, N. A. Benatar, J. Goodey, M. J. O'Doherty, and M. A. Smith, PET-MR image fusion in soft tissue sarcoma: accuracy, reliability and practicality of interactive point-based and automated mutual information techniques, *Eur J Nucl Med Mol Imaging* **30**, 54-62, 2003.

[76] G. J. Forster, C. Laumann, O. Nickel, P. Kann, O. Rieker, and P. Bartenstein, SPET/CT image co-registration in the abdomen with a simple and cost- effective tool, *Eur J Nucl Med Mol Imaging* **30**, 32-39, 2003.

[77] J. N. Yu, F. H. Fahey, B. A. Harkness, H. D. Gage, C. G. Eades, and J. W. Keyes, Jr., Evaluation of emission-transmission registration in thoracic PET, *J Nucl Med* **35**, 1777-1780, 1994.

[78] J. N. Yu, F. H. Fahey, H. D. Gage, C. G. Eades, B. A. Harkness, C. A. Pelizzari, and J. W. Keyes, Jr., Intermodality, retrospective image registration in the thorax, *J Nucl Med* **36**, 2333-2338, 1995.

[79] A. M. Scott, H. A. Macapinlac, C. R. Divgi, J. J. Zhang, H. Kalaigian, K. Pentlow, S. Hilton, M. C. Graham, G. Sgouros, C. Pelizzari, and et al., Clinical validation of SPECT and CT/MRI image registration in radiolabeled monoclonal antibody studies of colorectal carcinoma, *J Nucl Med* **35**, 1976-1984, 1994.

[80] A. M. Scott, H. Macapinlac, Z. Zhang, F. Daghighian, N. Montemayor, H. Kalaigian, G. Sgouros, M. C. Graham, K. Kolbert, S. D. J. Yeh, E. Lai, S. J. Goldsmith, and S. M. Larson, Image registration of SPECT and CT images using external fiduciary band and three-dimensional surface fitting in metastatic thyroid cancer, *J Nucl Med* **36**, 100-103, 1995.

[81] C. Perault, C. Schvartz, H. Wampach, J. C. Liehn, and M. J. Delisle, Thoracic and abdominal SPECT-CT image fusion without external markers in endocrine carcinomas. The Group of Thyroid Tumoral Pathology of Champagne-Ardenne, *J Nucl Med* **38**, 1234-1242, 1997.

[82] M. C. Gilardi, G. Rizzo, A. Savi, and F. Fazio, Registration of multi-modal biomedical images of the heart, *Q J Nucl Med* **40**, 142-150, 1996.

[83] J. Declerck, J. Feldmar, M. L. Goris, and F. Betting, Automatic registration and alignment on a template of cardiac stress and rest reoriented SPECT images, *IEEE Trans Med Imaging* **16**, 727-737, 1997.

[84] M. C. Gilardi, G. Rizzo, A. Savi, C. Landoni, V. Bettinardi, C. Rossetti, G. Striano, and F. Fazio, Correlation of SPECT and PET cardiac images by a surface matching registration technique, *Comput Med Imaging Graph* **22**, 391-398, 1998.

[85] J. Yang, S. C. Huang, M. Mega, K. P. Lin, A. W. Toga, G. W. Small, and M. E. Phelps, Investigation of partial volume correction methods for brain FDG PET studies, *IEEE Trans Nucl Sci* **43**, 3322-3327, 1996.

[86] H. W. Muller-Gartner, J. M. Links, J. L. Prince, R. N. Bryan, E. McVeigh, J. P. Leal, C. Davatzikos, and J. J. Frost, Measurement of radiotracer concentration in brain gray matter using positron emission tomography: MRI-based correction for partial volume effects, *J Cereb Blood Flow Metab* **12**, 571-583, 1992.

[87] D. Strul and B. Bendriem, Robustness of anatomically guided pixel-by-pixel algorithms for partial volume effect correction in positron emission tomography, *J Cereb Blood Flow Metab* **19**, 547-559, 1999.

[88] T. G. Turkington, T. R. DeGrado, M. W. Hanson, and R. E. Coleman, Alignment of dynamic cardiac PET images for correction of motion, *IEEE Trans Nucl Sci* **44**, 235-242, 1997.

[89] G. J. Klein, R. W. Reutter, and R. H. Huesman, Four-dimensional affine registration models for respiratory-gated PET, *IEEE Trans Nucl Sci* **48**, 756-760, 2001.

[90] N. Matsumoto, D. S. Berman, P. B. Kavanagh, J. Gerlach, S. W. Hayes, H. C. Lewin, J. D. Friedman, and G. Germano, Quantitative assessment of motion artifacts and validation of a new motion-correction program for myocardial perfusion SPECT, *J Nucl Med* **42**, 687-694, 2001.

[91] B. F. Hutton, A. Z. Kyme, Y. H. Lau, D. W. Skerrett, and R. R. Fulton, A hybrid 3-D reconstruction/registration algorithm for correction of head motion in emission tomography, *IEEE Trans Nucl Sci* **49**, 188-194, 2002.

[92] A. Z. Kyme, B. F. Hutton, R. L. Hatton, D. W. Skerrett, and L. R. Barnden, Practical aspects of a data-driven motion correction approach for brain SPECT, *IEEE Trans Med Imag* **22**, 722-729, 2003.

[93] B. Westermann and R. Hauser, Online head motion tracking applied to the patient registration problem, *Comput Aided Surg* **5**, 137-147, 2000.

[94] B. Ardekani, M. Braun, B. Hutton, I. Kanno, and H. Iida, Minimum cross-entropy reconstruction of PET images using prior anatomical information, *Phys Med Biol* **41**, 2497-2517, 1996.

[95] S. Som, B. F. Hutton, and M. Braun, Properties of minimum cross-entropy reconstruction with anatomical prior images, *IEEE Trans Nucl Sci* **45**, 3014-3021, 1998.

[96] P. Calvini, P. Vitali, F. Nobili, and G. Rodriquez, Enhancement of SPECT reconstructions by means of coregistered MR data, *IEEE Trans Nucl Sci* **48**, 750-755, 2001.

[97] G. W. Goerres, E. Kamel, T. N. Heidelberg, M. R. Schwitter, C. Burger, and G. K. von Schulthess, PET-CT image co-registration in the thorax: influence of respiration, *Eur J Nucl Med Mol Imaging* **29**, 351-360, 2002.

[98] H. M. Hudson and R. S. Larkin, Accelerated image reconstruction using ordered subsets of projection data, *IEEE Trans Med Imag* **13**, 601-609, 1994.

[99] L. G. Brown, A survey of image registration techniques, *ACM Computing Surveys (CSUR)* **24**, 325-376, 1992.

[100] H. Lester and S. R. Arridge, A survey of hierarchical non-linear medical image registration, *Pattern Recognition* **32**, 129-149, 1999.

[101] D. J. Hawkes, Algorithms for radiological image registration and their clinical application, *J Anat* **193**, 347-361, 1998.

[102] B. F. Hutton and M. Braun, Software for image registration: algorithms, accuracy, efficacy, *Semin Nucl Med* **33**, 180-92, 2003.

[103] W. L. Nowinski and A. Thirunavuukarasuu, Atlas-assisted localization analysis of functional images, *Med Image Anal* **5**, 207-220, 2001.

[104] B. A. Vogt, S. Derbyshire, and A. K. Jones, Pain processing in four regions of human cingulate cortex localized with co-registered PET and MR imaging, *Eur J Neurosci* **8**, 1461-1473, 1996.

[105] K. J. Friston, A. P. Holmes, K. J. Worsley, J.-B. Poline, C. D. Frith, and R. S. J. Frackowiak, Statistical parametric maps in functional imaging: a general linear approach, *Hum Brain Map* **2**, 189-210, 1995.

[106] J. Ashburner and K. Friston, Multimodal Image Coregistration and Partitioning--A Unified Framework, *Neuro Image* **6**, 209-217, 1997.

[107] J. Ashburner, P. Neelin, D. L. Collins, A. Evans, and K. Friston, Incorporating prior knowledge into image registration, *Neuroimage* **6**, 344-352, 1997.

[108] K. J. Friston and W. Penny, Posterior probability maps and SPMs, *Neuro Image* **19**, 1240-1249, 2003.

[109] K. J. Van Laere, J. Warwick, J. Versijpt, I. Goethals, K. Audenaert, B. Van Heerden, and R. Dierckx, Analysis of clinical brain SPECT data based on anatomic standardization and reference to normal data: an ROC-based comparison of visual, semiquantitative, and voxel-based methods, *J Nucl Med* **43**, 458-469, 2002.

[110] E. H. W. Meijering, W. J. Niessen, and M. A. Viergever, Retrospective Motion Correction in Digital Subtraction Angiography: A Review, *IEEE Trans. Med. Imag.* **18**, 2-21, 1999.

[111] K. J. Lee and D. C. Barber, Use of forward projection to correct patient motion during SPECT imaging, *Phys Med Biol* **43**, 171-187, 1998.

[112] K. A. Langmack, Portal imaging, *Br J Radiol* **74**, 789-804, 2001.

[113] P. Remeijer, E. Geerlof, L. Ploeger, K. Gilhuijs, M. van Herk, and J. V. Lebesque, 3-D portal image analysis in clinical practice: an evaluation of 2-D and 3-D analysis techniques as applied to 30 prostate cancer patients, *Int J Radiat Oncol Biol Phys* **46**, 1281-1290, 2000.

[114] D. Plattard, M. Soret, J. Troccaz, P. Vassal, J. Y. Giraud, G. Champleboux, X. Artignan, and M. Bolla, Patient set-up using portal images: 2D/2D image registration using mutual information, *Comput Aided Surg* **5**, 246-262, 2000.

[115] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. G. Hill, and D. J. hawkes, A comparison of similarity measures for use in 2-D-3-D medical image registration, *IEEE Trans Med Imag* **17**, 586-595, 1998.

[116] K. S. Arun, T. S. Huang, and S. D. Blostein, Least-Squares fitting of two 3D point sets, *IEEE Trans Patt Anal Mach Intell* **9**, 696-700, 1987.

[117] P. A. Viola, "Alignment by Maximization of Mutual Information," Massachusetts Institute of Technology, Boston, MA, USA, Thesis Date 1995.

[118] P. Thevenaz, U. E. Ruttimann, and M. Unser, A pyramid approach to subpixel registration based on intensity, *IEEE Trans Imag Proc* **7**, 27-41, 1998.

[119] X. Pennec and J. P. Thirion, Validation of 3-D registration methods based on points and frames, *Fifth International Conference on Computer Vision Proceedings*, pp.557-562, 1995.

[120] D. L. Hill, D. J. Hawkes, J. E. Crossman, M. J. Gleeson, T. C. Cox, E. E. Bracey, A. J. Strong, and P. Graves, Registration of MR and CT images for skull base surgery using point- like anatomical features, *Br J Radiol* **64**, 1030-1035, 1991.

[121] C. R. Maurer, Jr., J. M. Fitzpatrick, M. Y. Wang, R. L. Galloway, Jr., R. J. Maciunas, and G. S. Allen, Registration of head volume images using implantable fiducial markers, *IEEE Trans Med Imaging* **16**, 447-462, 1997.

[122] J. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, C. R. Maurer, Jr., R. M. Kessler, R. J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. A. van den Elsen, S. Napel, T. S. Sumanaweera, B. Harkness, P. F. Hemler, D. L. Hill, D. J. Hawkes, C. Studholme, J. B. Maintz, M. A. Viergever, G. Malandain, R. P. Woods, and et al., Comparison and evaluation of retrospective intermodality brain image registration techniques, *J Comput Assist Tomogr* **21**, 554-566, 1997.

[123] L. Barnden, R. Kwiatek, Y. Lau, B. Hutton, L. Thurfjell, K. Pile, and C. Rowe, Validation of fully automatic brain SPET to MR co-registration, *Eur J Nucl Med* **27**, 147-154, 2000.

[124] P. H. Schonemann, A generalized solution of the orthogonal procrustes problem, *Psychometrika* **31**, 1-10, 1966.

[125] J. M. Fitzpatrick and J. B. West, The distribution of target registration error in rigid-body point-based registration, *IEEE Trans Med Imaging* **20**, 917-927, 2001.

[126] M. A. Audette, F. P. Ferrie, and T. M. Peters, An algorithmic overview of surface registration techniques for medical imaging, *Med Image Anal* **4**, 201-217, 2000.

[127] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, vol., 2nd ed. Cambridge: Cambridge University Press, 1992.

[128] I. D. Grachev, D. Berdichevsky, S. L. Rauch, S. Heckers, D. N. Kennedy, V. S. Caviness, and N. M. Alpert, A method for assessing the accuracy of intersubject registration of the human brain using anatomic landmarks, *Neuroimage* **9**, 250-268, 1999.

[129] A. Panigrahy, S. D. Caruthers, J. Krejza, P. D. Barnes, S. G. Faddoul, L. A. Sleeper, and E. R. Melhem, Registration of three-dimensional MR and CT studies of the cervical spine, *AJNR Am J Neuroradiol* **21**, 282-289, 2000.

[130] J. C. Liehn, A. Loboguerrero, C. Perault, and L. Demange, Superimposition of computed tomography and single photon emission tomography immunoscintigraphic images in the pelvis: validation in patients with colorectal or ovarian carcinoma recurrence, *Eur J Nucl Med* **19**, 186-194, 1992.

[131] B. A. Ardekani, M. Braun, B. F. Hutton, I. Kanno, and H. Iida, A fully automatic multimodality image registration algorithm, *J Comput Assist Tomogr* **19**, 615-623, 1995.

[132] J.-P. Thirion, Extremal points: definition and application to 3D image registration, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR '94*, pp.587-592, 1994.

[133] L. Y. Hsu, M. H. Loew, and J. Ostuni, Automated registration of brain images using edge and surface features, *IEEE Eng Med Biol Mag* **18**, 40-47, 1999.

[134] M. Kass, A. Witkin, and D. Terzopoulos, Snakes: Active contour models, *Int J Comput Vision* **1**, 321-331, 1988.

[135] C. C. Poon and M. Braun, Image segmentation by a deformable contour model incorporating region analysis, *Phys Med Biol* **42**, 1833-1841, 1997.

[136] I. Kapouleas, A. Alavi, W. M. Alves, R. E. Gur, and D. W. Weiss, Registration of three-dimensional MR and PET images of the human brain without markers, *Radiology* **181**, 731-739, 1991.

[137] H. Li, B. S. Manjunath, and S. K. Mitra, Registration of 3D multimodality brain images by curve matching, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR '94 Proceedings*, pp.1744-1748, 1994.

[138] Y. Ge, J. M. Fitzpatrick, J. R. Votaw, S. Gadamsetty, R. J. Maciunas, R. M. Kessler, and R. A. Margolin, Retrospective registration of PET and MR brain images: an algorithm and its stereotactic validation, *J Comput Assist Tomogr* **18**, 800-810, 1994.

[139] J. B. A. Maintz, P. A. van den Elsen, and M. A. Viergever, Comparison of edge-based and ridge-based registration of CT and MR brain images, *Med Image Anal* **1**, 151-161, 1996.

[140] J.-P. Thirion, Direct extraction of boundaries from computed tomography scans, *IEEE Trans Med Imag* **13**, 322-328, 1994.

[141] J.-P. Thirion, The extremal mesh and the understanding of 3D surfaces, I*EEE Workshop on Biomedical Image Analysis Proceedings*, pp.3-12, 1994.

[142] A. Gueziec, Surface Representation with Deformable Splines: Using Decoupled Variables, *IEEE Computational Science and Engineering* **2**, 69-80, 1995.

[143] A. P. Gueziec, X. Pennec, and N. Ayache, Medical image registration using geometric hashing, *IEEE Computational Science and Engineering* **4**, 29-41, 1997.

[144] C. Davatzikos, J. L. Prince, and R. N. Bryan, Image registration based on boundary mapping, *IEEE Trans Med Imag* **15**, 112-115, 1996.

[145] C. A. Pelizzari, G. T. Chen, D. R. Spelbring, R. R. Weichselbaum, and C. T. Chen, Accurate three-dimensional registration of CT, PET, and/or MR images of the brain, *J Comput Assist Tomogr* **13**, 20-26, 1989.

[146] D. N. Levin, C. A. Pelizzari, G. T. Chen, C. T. Chen, and M. D. Cooper, Retrospective geometric correlation of MR, CT, and PET images, *Radiology* **169**, 817-823, 1988.

[147] B. Wang, C. Toro, T. A. Zeffiro, and M. Hallett, Head surface digitization and registration: a method for mapping positions on the head onto magnetic resonance images, *Brain Topogr* **6**, 185-192, 1994.

[148] P. J. Besl and N. D. McKay, A method for registration of 3-D shapes, *IEEE Trans Pattern Anal Mach Intell* **14**, 239-256, 1992.

[149] Z. Zhang, Iterative Point Matching for Registration of Free-Form Curves and Surfaces, *Int J Comput Vision* **13**, 119-152, 1994.

[150] J. Feldmar and N. Ayache, Locally affine registration of free-form surfaces, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR '94 Proceedings*, pp.496-501, 1994.

[151] J. Feldmar, G. Malandain, J. Declerck, and N. Ayache, Extension of the ICP algorithm to non-rigid intensity-based registration of 3D volumes, *Workshop on Mathematical Methods in Biomedical Image Analysis Proceedings*, pp.84-93, 1996.

[152] C. R. Maurer, Jr., D. L. Hill, A. J. Martin, H. Liu, M. McCue, D. Rueckert, D. Lloret, W. A. Hall, R. E. Maxwell, D. J. Hawkes, and C. L. Truwit, Investigation of intraoperative brain deformation using a 1.5-T interventional MR system: preliminary results, *IEEE Trans Med Imaging* **17**, 817-825, 1998.

[153] G. Borgefors, Distance Transformations in Digital Images, *Comput Vision, Graphics, Image Processing* **34**, 344-371, 1986.

[154] G. Borgefors, Hierarchical chamfer matching: a parametric edge matching algorithm, *IEEE Trans Patt Anal Mach Intell* **10**, 849-865, 1988.

[155] M. van Herk, J. C. de Munck, J. V. Lebesque, S. Muller, C. Rasch, and A. Touw, Automatic registration of pelvic computed tomography data and magnetic resonance scans including a full circle method for quantitative accuracy evaluation, *Med Phys* **25**, 2054-2067, 1998.

[156] V. S. Khoo, E. J. Adams, F. Saran, J. L. Bedford, J. R. Perks, A. P. Warrington, and M. Brada, A Comparison of clinical target volumes determined by CT and MRI for the radiotherapy planning of base of skull meningiomas, *Int J Radiat Oncol Biol Phys* **46**, 1309-1317, 2000.

[157] P. A. van den Elsen, J. B. A. Maintz, E. J. D. Pol, and M. A. Viergever, Image fusion using geometrical features, *SPIE Visualization in Biomedical Computing Proceedings*, pp.172-186, 1992.

[158] J. F. Mangin, V. Frouin, I. Bloch, B. Bendriem, and J. Lopez-Krahe, Fast nonsupervised 3D registration of PET and MR images of the brain, *J Cereb Blood Flow Metab* **14**, 749-762, 1994.

[159] J. C. de Munck, F. C. Verster, E. A. Dubois, J. B. Habraken, B. Boltjes, J. J. Claus, and M. van Herk, Registration of MR and SPECT without using external fiducial markers, *Phys Med Biol* **43**, 1255-1269, 1998.

[160] J. Cai, J. C. Chu, D. Recine, M. Sharma, C. Nguyen, R. Rodebaugh, V. A. Saxena, and A. Ali, CT and PET lung image registration and fusion in radiotherapy treatment planning using the chamfer-matching method, *Int J Radiat Oncol Biol Phys* **43**, 883-891, 1999.

[161] N. M. Alpert, J. F. Bradshaw, D. Kennedy, and J. A. Correia, The principal axes transformation--a method for image registration, *J Nucl Med* **31**, 1717-1722, 1990.

[162] H. Rusinek, W. H. Tsui, A. V. Levy, M. E. Noz, and M. J. de Leon, Principal axes and surface fitting methods for three-dimensional image registration, *J Nucl Med* **34**, 2019-2024, 1993.

[163] A. P. Dhawan, L. K. Arata, A. V. Levy, and J. Mantil, Iterative Principal Axes Registration method for analysis of MR-PET brain images, *IEEE Trans Biomed Eng* **42**, 1079-1087, 1995.

[164] C. K. Hoh, M. Dahlbom, G. Harris, Y. Choi, R. A. Hawkins, M. E. Phelps, and J. Maddahi, Automated iterative three-dimensional registration of positron emission tomography images, *J Nucl Med* **34**, 2009-2018, 1993.

[165] S. Eberl, D. Kacher, R. Fulton, A. Ryan, B. Hutton, and M. Fulham, Automated interstudy image registartion technique for SPECT and PET, *J Nucl Med* **37**, 137-145, 1996.

[166] J. V. Hajnal, N. Saeed, E. J. Soar, A. Oatridge, I. R. Young, and G. M. Bydder, A registration and interpolation procedure for subvoxel matching of serially acquired MR images, *J Comput Assist Tomogr* **19**, 289-296, 1995.

[167] I. G. Zubal, S. S. Spencer, K. Imam, J. Seibyl, E. O. Smith, G. Wisniewski, and P. B. Hoffer, Difference images calculated from ictal and interictal technetium-99m- HMPAO SPECT scans of epilepsy, *J Nucl Med* **36**, 684-689, 1995.

[168] P. Vera, A. Kaminska, C. Cieuta, A. Hollo, J. L. Stievenart, I. Gardin, D. Ville, J. F. Mangin, P. Plouin, O. Dulac, and C. Chiron, Use of subtraction ictal SPECT co-registered to MRI for optimizing the localization of seizure foci in children, *J Nucl Med* **40**, 786-792, 1999.

[169] P. A. van den Elsen, J. B. A. Maintz, E.-J. D. Pol, and M. A. Viergever, Automatic registration of CT and MR brain images using correlation of geometrical features, *IEEE Trans Med Imag* **14**, 384-396, 1995.

[170] L. Junck, J. G. Moen, G. D. Hutchins, M. B. Brown, and D. E. Kuhl, Correlation methods for the centering, rotation and alignment of functional brain images., *J Nucl Med* **31**, 1220-1276, 1990.

[171] D. L. Collins, P. Neelin, T. M. Peters, and A. C. Evans, Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space, *J Comput Assist Tomogr* **18**, 192-205, 1994.

[172] R. P. Woods, J. C. Mazziotta, and S. R. Cherry, MRI-PET registration with automated algorithm, *J Comput Assist Tomogr* **17**, 536-546, 1993.

[173] D. L. G. Hill, C. Studholme, and D. J. Hawkes, Voxel similarity measures for automated image registration, *Visualization in Biomedical Computing, SPIE Press*, pp.205-216, 1994.

[174] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "Automated 3D MR and PET brain image registration," in *Computer Asisted Radiology*, vol., H. U. Lemke, K. Imamura, C. C. Jaffe, and M. W. Vannier, Eds., Berlin: Springer-Verlag, 1995, pp.248-253.

[175] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "Multiresolution voxel similarity measures for MR-PET registration," in *Information processing in Medical Imaging 14th International Conference*, vol., Y. Bizais, C. Barillot, and R. D. Paola, Eds., Dordrecht: Kluwer Academic, 1995, pp.287-298.

[176] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, and P. Suetens, "Automated multimodality image registration using information theory," in *Information Processing in Medical Imaging*, vol., Y. Bizais and C. Barillot, Eds., Dordrecht: Kluwer Academic Publishers, 1995, pp.263-274.

[177] P. Viola and W. M. Wells III, "Alignment by maximization of mutual information," in *Int. Conf. on Computer Vision*, vol., E. Grimson, S. Shafer, A. Blake, and K. Sugihara, Eds., IEEE Computer Society Press, Los Alamitos, CA., 1995, pp.16-23.

[178] W. M. Wells, 3rd, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, Multi-modal volume registration by maximization of mutual information, *Med Image Anal* **1**, 35-51, 1996.

[179] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, Multimodality image registration by maximization of mutual information, *IEEE Trans Med Imaging* **16**, 187-198, 1997.

[180] C. Studholme, D. L. G. Hill, and D. J. Hawkes, An overlap invariant entropy measure of 3D medical image alignment, *Pattern Recognition* **32**, 71-86, 1999.

[181] A. Roche, G. Malandain, X. Pennec, and N. Ayache, "The correlation ratio as a new similarity measure for multimodal image registration," in *First International Conference on Medical Image Computing and Computer-Assisted Intervention MICCAI'98*, vol., *Lecture Notes in Computer Science 1496*, W. M. Wells, A. Colchester, and S. Delp, Eds., Cambridge, MA, USA: Berlin: Springer-Verlag, 1998, pp.1115-1124.

[182] Y. H. Lau, M. Braun, and B. F. Hutton, Non-rigid image registration using a median-filtered coarse-to-fine displacement field and a symmetric correlation ratio, *Phys Med Biol* **46**, 1297-319, 2001.

[183] J. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, C. R. Maurer, Jr., R. M. Kessler, and R. J. Maciunas, Retrospective intermodality registration techniques for images of the head: surface-based versus volume-based, *IEEE Trans Med Imaging* **18**, 144-150, 1999.

[184] G. C. Kagadis, K. K. Delibasis, G. K. Matsopoulos, N. A. Mouravliansky, P. A. Asvestas, and G. C. Nikiforidis, A comparative study of surface- and volume-based techniques for the automatic registration between CT and SPECT brain images, *Med Phys* **29**, 201-213, 2002.

[185] N. Ritter, R. Owens, J. Cooper, R. H. Eikelboom, and P. P. van Saarloos, Registration of stereo and temporal images of the retina, *IEEE Trans Med Imaging* **18**, 404-418, 1999.

[186] C. R. Meyer, J. L. Boes, B. Kim, P. H. Bland, K. R. Zasadny, P. V. Kison, K. Koral, K. A. Frey, and R. L. Wahl, Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations, *Med Image Anal* **1**, 195-206, 1997.

[187] D. L. G. Hill, C. R. J. Maurer, C. Studholme, J. M. Fitzpatrick, and D. J. Hawkes, Correcting scaling errors in tomographic images using a nine degree of freedom registration algorithm, *J Comput Assist Tomogr* **22**, 317-323, 1998.

[188] C. Studholme, J. A. Little, G. P. Penney, D. L. G. Hill, and D. J. Hawkes, "Automated multimodality registartion using the full affine transformation: application to MR and CT guided skull base surgery," in *Visualization in Biomedical Computing*, vol. 1131, *Lecture Notes in Computer Science*, K. H. Hohne and R. Kikinis, Eds., Springer-Verlag, Berlin, 1996, pp.601-606.

[189] P. Viola and W. M. Wells III, Alignment by maximization of mutual information, *Int J Comput Vision* **24**, 137-154, 1997.

[190] J. F. Krucker, C. R. Meyer, G. L. LeCarpentier, J. B. Fowlkes, and P. L. Carson, 3D spatial compounding of ultrasound images using image-based nonrigid registration, *Ultrasound Med Biol* **26**, 1475-1488, 2000.

[191] L. Hadjiiski, H. P. Chan, B. Sahiner, N. Petrick, and M. A. Helvie, Automated registration of breast lesions in temporal pairs of mammograms for interval change analysis--local affine transformation for improved localization, *Med Phys* **28**, 1070-1079, 2001.

[192] M. Jenkinson and S. Smith, A global optimisation method for robust affine registration of brain images, *Med Image Anal* **5**, 143-156, 2001.

[193] U. Kjems, S. C. Strother, J. Anderson, I. Law, and L. K. Hansen, Enhancing the multivariate signal of [^{15}O] water PET studies with a new nonlinear neuroanatomical registration algorithm [MRI application], *IEEE Trans Med Imag* **18**, 306-319, 1999.

[194] P. J. Edwards, A. P. King, D. J. Hawkes, O. Fleig, C. R. Maurer, Jr., D. L. Hill, M. R. Fenlon, D. A. de Cunha, R. P. Gaston, S. Chandra, J. Mannss, A. J. Strong, M. J. Gleeson, and T. C. Cox, Stereo augmented reality in the surgical microscope, *Stud Health Technol Inform* **62**, 102-108, 1999.

[195] P. Kotsas, S. Malasiotis, M. Strintzis, D. W. Piraino, and J. F. Cornhill, A fast and accurate method for registration of MR images of the head, *Int J Med Inf* **52**, 167-182, 1998.

[196] M. E. Noz, G. Q. Maguire, Jr., M. P. Zeleznik, E. L. Kramer, F. Mahmoud, and J. Crafoord, A versatile functional-anatomic image fusion method for volume data sets, *J Med Syst* **25**, 297-307, 2001.

[197] S. Kimiaei, M. Noz, E. Jonsson, J. Crafoord, and G. Q. Maguire, Jr., Evaluation of polynomial image deformation using anatomical landmarks for matching of 3D-abdominal MR-images and for atlas construction, *IEEE Trans Nucl Sci* **46**, 1110-1113, 1999.

[198] F. L. Bookstein, Principal warps: Thin-plate splines and the decomposition of deformations, *IEEE Trans Patt Anal Mach Intell* **11**, 567-585, 1989.

[199] R. Bajcsy and S. Kovacic, Multiresolution elastic matching, *Comput Vision, Graphics Image Proc* **46**, 1-12, 1989.

[200] H. D. Malcolm, K. Alireza, P. F. Duane, and E. H. Steven, A physics-based coordinate transformation for 3-D image matching, *IEEE Trans Med Imag* **16**, 317-328, 1997.

[201] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, 3D Brain mapping using deformable neuroanatomy, *Phys Med Biol* **39**, 609-618, 1994.

[202] M. Bro-Nielsen and C. Gramkow, "Fast fluid registration of medical images," in *Proc. in Visualization in Biomedical Computing*, vol. 1131, *Lecture Notes in Computer Science*, K. H. Hohne and R. Kikinis, Eds., Hamburg: Springer-Verlag, 1996, pp.267-276.

[203] P. J. Slomka, D. Dey, C. Przetak, U. E. Aladl, and R. P. Baum, Automated 3-dimensional registration of stand-alone (18)F-FDG whole-body PET with CT, *J Nucl Med* **44**, 1156-1167, 2003.

[204] J. A. Little, D. L. G. Hill, and D. J. Hawkes, Deformations incorporating rigid structures, *computer Vision Image Understanding* **66**, 223-232, 1997.

[205] K. S. Rohr, H.S.; Sprengel, R.; Buzug, T.M.; Weese, J.; Kuhn, M.H., Landmark-based elastic registration using approximating thin-plate splines, *IEEE Trans Med Imag* **20**, 526-534, 2001.

[206] H. J. Johnson and G. E. Christensen, Consistent landmark and intensity-based image registration, *IEEE Trans Med Imag* **21**, 450-461, 2002.

[207] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, Nonrigid registration using free-form deformations: application to breast MR images, *IEEE Trans Med Imag* **18**, 712-721, 1999.

[208] B. Likar and F. Pernus, Registration of serial transverse sections of muscle fibers, *Cytometry* **37**, 93-106, 1999.

[209] C. Studholme, R. T. Constable, and J. S. Duncan, Accurate alignment of functional EPI data to anatomical MRI using a physics-based distortion model, *IEEE Trans Med Imaging* **19**, 1115-1127, 2000.

[210] N. Arad and D. Reisfeld, Image warping using few anchor points and radial functions, *Comp Graphics Forum* **14**, 35-46, 1995.

[211] T. Gaens, F. Maes, D. Vandermeulen, and P. Suetens, "Non-rigid multimodal image registration using mutual information," in *First Int. Conf. on Medical Image Computing and Computer-Assisted Intervention MICCAI'98*, vol. 1496, *Lecture Notes in Computer Science*, W. M. Wells, A. Colchester, and S. Delp, Eds., Berlin: Springer-Verlag, 1998, pp.1099-1106.

[212] E. Maeland, On the comparison of interpolation methods, *IEEE Trans Med Imag* **7**, 213-217, 1988.

[213] T. M. Lehmann, C. Gonner, and K. Spitzer, Addendum: B-spline interpolation in medical image processing, *IEEE Trans Med Imag* **20**, 660-665, 2001.

[214] P. Thevenaz, T. Blu, and M. Unser, Interpolation revisited [medical images application], *IEEE Trans Med Imag* **19**, 739-758, 2000.

[215] E. H. W. Meijering, W. J. Niessen, and M. A. Viegever, Quantitative evaluation of convolution-based methods for medical image interpolation, *Med Image Anal* **5**, 111-126, 2001.

[216] J. A. Parker, R. V. Kenyon, and D. E. Troxel, Comparison of interpolating methods for image resampling, *IEEE Trans Med Imag* **MI-2**, 31-39, 1983.

[217] C. R. Appledorn, A new approach to the interpolation of sampled data, *IEEE Trans Med Imag* **15**, 369-376, 1996.

[218] P. Thevenaz and M. Unser, Optimization of mutual information for multiresolution image registration, *IEEE Trans Image Processing* **9**, 2083-2099, 2000.

[219] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, Interpolation artefacts in mutual information-based image registration, *computer Vision Image Understanding* **77**, 211-232, 2000.

[220] A. Collignon, "Multi-modality medical image registration by maximization of mutual information," Catholic University of Leuven, Leuven, Belgium, Thesis Date 1998.

[221] F. Maes, "Segmentation and registration of multimodal medical images: From theory, implementation and validation to a useful tool in clinical practice," Catholic University of Leuven, Leuven, Belgium, Thesis Date 1998.

[222] G. J. Grevera and J. K. Udupa, Shape-based interpolation of multidimensional grey-level images, *IEEE Trans Med Imag* **15**, 881-892, 1996.

[223] K. Kanatani, Analysis of 3-D rotation fitting, *IEEE Trans Patt Anal Mach Intell* **16**, 543-549, 1994.

[224] J. M. Fitzpatrick, J. B. West, and C. R. Maurer, Jr., Predicting error in rigid-body point-based registration, *IEEE Trans Med Imaging* **17**, 694-702, 1998.

[225] F. Maes, D. Vandermeulen, and P. Suetens, Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information, *Med Image Anal* **3**, 373-386, 1999.

[226] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, Image registration by maximization of combined mutual information and gradient information, *IEEE Trans Med Imag* **19**, 809-814, 2000.

[227] B. H. Brinkmann, T. J. O'Brien, S. Aharon, M. K. O'Connor, B. P. Mullan, D. P. Hanson, and R. A. Robb, Quantitative and clinical analysis of SPECT image registration for epilepsy studies, *J Nucl Med* **40**, 1098-1105, 1999.

[228] C. Nikou, F. Heitz, J. P. Armspach, I. J. Namer, and D. Grucker, Registration of MR/MR and MR/SPECT brain images by fast stochastic optimization of robust voxel similarity measures, *Neuroimage* **8**, 30-43, 1998.

[229] G. K. Matsopoulos, N. A. Mouravliansky, K. K. Delibasis, and K. S. Nikita, Automatic retinal image registration scheme using global optimization techniques, *IEEE Trans Inf Technol Biomed* **3**, 47-60, 1999.

[230] H. Renjie and P. A. Narayana, Global optimization of mutual information: application to three-dimensional retrospective registration of magnetic resonance images, *Computerized Medical Imaging and Garphics* **26**, 277-292, 2002.

[231] J. M. Rouet, J. J. Jacq, and C. Roux, Genetic algorithms for a robust 3-D MR-CT registration, *IEEE Trans Inf Technol Biomed* **4**, 126-136, 2000.

[232] C. Studholme, D. L. Hill, and D. J. Hawkes, Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures, *Med Phys* **24**, 25-35, 1997.

[233] T. Pfluger, C. Vollmar, A. Wismuller, S. Dresel, F. Berger, P. Suntheim, G. Leinsinger, and K. Hahn, Quantitative comparison of automatic and interactive methods for MRI- SPECT image registration of the brain based on 3-dimensional calculation of error, *J Nucl Med* **41**, 1823-1829, 2000.

[234] A. R. Peters, S. H. Muller, J. C. de Munck, and M. van Herk, The accuracy of image registration for the brain and the nasopharynx using external anatomical landmarks, *Phys Med Biol* **45**, 2403-2416, 2000.

[235] O. Sipila, P. Nikkinen, H. Pohjonen, V. P. Poutanen, A. Visa, S. Savolainen, T. Katila, and K. Liewendahl, Accuracy of a registration procedure for brain SPET and MRI: phantom and simulation studies, *Nucl Med Commun* **18**, 517-526, 1997.

[236] J. B. West, J. M. Fitzpatrick, S. A. Toms, C. R. Maurer, Jr., and R. J. Maciunas, Fiducial point placement and the accuracy of point-based, rigid body registration, *Neurosurgery* **48**, 810-816; discussion 816-817, 2001.

[237] J. M. Fitzpatrick, D. L. G. Hill, Y. Shyr, J. West, C. Studholme, and C. R. Maurer, Jr., Visual assessment of the accuracy of retrospective registration of MR and CT images of the brain, *IEEE Trans Med Imag* **17**, 571-585, 1998.

[238] W. Zylka and J. Sabczynski, Effect of localization devices and registration methods on the accuracy of stereotactic frame systems predicted by the Gaussian approach, *Comput Aided Surg* **4**, 77-86, 1999.

[239] R. P. Woods, S. T. Grafton, C. J. Holmes, S. R. Cherry, and J. C. Mazziotta, Automated image registration: I. General methods and intrasubject, intramodality validation, *J Comput Assist Tomogr* **22**, 139-152, 1998.

[240] R. P. Woods, S. T. Grafton, J. D. Watson, N. L. Sicotte, and J. C. Mazziotta, Automated image registration: II. Intersubject validation of linear and nonlinear models, *J Comput Assist Tomogr* **22**, 153-165, 1998.

[241] J. Talairach and P. Tournoux, *Co-planar stereotaxic atlas of the human brain. 3-dimensional proportional system: An approach to cerebral imaging*, New York: Thieme Medical, 1988.

[242] M. Holden, D. L. G. Hill, E. R. E. Denton, J. M. Jarosz, T. C. S. Cox, T. Rohlfing, J. Goodey, and D. J. Hawkes, Voxel similarity measures for 3-D serial MR brain image registration, *IEEE Trans Med Imag* **19**, 94-102, 2000.

[243] Y. M. Zhu and S. M. Cochoff, Influence of implementation parameters on registration of MR and SPECT brain images by maximization of mutual information, *J Nucl Med* **43**, 160-166, 2002.

[244] C. Studholme, D. L. Hill, and D. J. Hawkes, Automated 3-D registration of MR and CT images of the head, *Med Image Anal* **1**, 163-175, 1996.

[245] D. J. Hawkes, D. L. G. Hill, E. D. Lehmann, G. P. Robinson, M. N. Maisey, and A. C. F. Colchester, "Preliminary work on the interpretation of SPECT images with the aid of registered MR images and an MR derived 3D neuro-anatomical atlas," in *3D Imaging in Medicine*, vol., K. H. Hohne, H. Fuchs, and S. M. Pizer, Eds., Berlin: Springer-Verlag, 1990, pp.242-251.

[246] K. Rehm, S. C. Strother, J. R. Anderson, K. A. Schaper, and D. A. Rottenberg, Display of merged multimodality brain images using interleaved pixels with independent color scales, *J Nucl Med* **35**, 1815-1821, 1994.

[247] M. Quarantelli, B. Alfano, M. Larobina, E. Tedeschi, A. Brunetti, E. M. Covelli, A. Ciarmiello, C. Mainolfi, and M. Salvatore, Frequency encoding for simultaneous display of multimodality images, *J Nucl Med* **40**, 442-447, 1999.

[248] J. C. Wong, C. Studholme, D. J. Hawkes, and M. N. Maisey, Evaluation of the limits of visual detection of image misregistration in a brain fluorine-18 fluorodeoxyglucose PET-MRI study, *Eur J Nucl Med* **24**, 642-650, 1997.

[249] G. Zubal, G. Gindi, M. Lee, C. Harrell, and E. Smith, High resolution anthropomorphic phantom for Monte Carlo analysis of internal radiation sources, *IEEE Symp Computer-Based Medical Systems*, pp.540-547, 1990.

[250] G. Zubal, C. R. Harrell, E. O. Smith, Z. Rattner, G. Gindi, and P. B. Hoffer, Computerized three-dimensional segmented human anatomy, *Med Phys* **21**, 299-302, 1994.

[251] V. M. Spitzer and D. G. Whitlock, The Visible Human Dataset: the anatomical platform for human simulation, *Anat Rec* **253**, 49-57, 1998.

[252] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans, Design and construction of a realistic digital brain phantom, *IEEE Trans Med Imaging* **17**, 463-468, 1998.

[253] L. Livieratos, S. R. Meikle, J. C. Matthews, D. L. Bailey, T. Jones, and P. Price, Assessment of coregistration of PET $H_2{}^{15}O$ images in the abdomen using digital phantoms, *IEEE Nuclear Science Symposium*, pp.1175-1178, 1997.

[254] K. Sjogreen, M. Ljungberg, K. Erlandsson, L. Floreby, and S. E. Strand, "Registartion of abdominal CT and SPECT images using Compton scatter data," in *Information Processing in Medical Imaging*, vol., J. S. Duncan and G. Gindi, Eds., Berlin, Germany: Springer-Verlag, 1997, pp.232-244.

[255] P. Neelin, J. Crossman, D. J. Hawkes, Y. Ma, and A. C. Evans, Validation of an MRI/PET landmark registration method using 3D simulated PET images and point simulations, *Comput Med Imaging Graph* **17**, 351-356, 1993.

[256] K.-P. Lin, S.-C. Huang, L. R. Baxter, and M. E. Phelps, A general technique for interstudy registration of multifunction and multimodality images, *IEEE Trans Nucl Sci* **41**, 2850-2855, 1994.

[257] S. J. Kiebel, J. Ashburner, J. B. Poline, and K. J. Friston, MRI and PET coregistration--a cross validation of statistical parametric mapping and automated image registration, *Neuroimage* **5**, 271-279, 1997.

[258] L. Thurfjell, Y. H. Lau, J. L. Andersson, and B. F. Hutton, Improved efficiency for MRI-SPET registration based on mutual information, *Eur J Nucl Med* **27**, 847-856, 2000.

[259] E. J. Hoffman, P. D. Cutler, T. M. Guerrero, W. M. Digby, and J. C. Mazziotta, Assessment of accuracy of PET ultizing a 3-D phantom to simulate the activity distribution of [^{18}F]fluorodeoxyglucose uptake in the human brain, *J Cereb Blood Flow Metab* **11**, 17-25, 1991.

[260] D. Dey, P. J. Slomka, L. J. Hahn, and R. Kloiber, Automatic three-dimensional multimodality registration using radionuclide transmission CT attenuation maps: a phantom study, *J Nucl Med* **40**, 448-455, 1999.

[261] H. R. Tang, J. K. Brown, A. J. Da Silva, K. K. Matthay, D. C. Price, J. P. Huberty, R. A. Hawkins, and B. H. Hasegawa, Implementation of a combined X-ray CT-scintillation camera imaging system for localizing and measuring radionuclide uptake: experiments in phantoms and patients, *IEEE Trans Nucl Sci* **46**, 551-557, 1999.

[262] W. K. Pratt, *Digital Image Processing*, vol., 2nd ed. New York: John Wiley & Sons, Inc., 1991.

[263] A. Roche, G. Malandain, and N. Ayache, Unifying Maximum Likelihood Approaches in Medical image registration, Inst. National de Recherche d'Informatique et d'Auto (INRIA), France, Report No. 3741, 1999.

[264] A. Roche, G. Malandain, and N. Ayache, Unifying Maximum Likelihood Approaches in Medical Image Registration, *International Journal of Imaging Systems and Technology: Special issue on 3D imaging* **11**, 71-80, 2000.

[265] A. Roche, G. Malandain, X. Pennec, and N. Ayache, Multimodal Image Registration by Maximization of the Correlation Ratio, Inst. National de Recherche d'Informatique et d'Auto (INRIA), France, Report No. 3378, 1998.

[266] C. E. Shannon, The mathematical theory of communication (part I and 2), *Bell Syst Tech J* **27**, 379-423, 623-656, 1948.

[267] C. E. Shannon, Communication In The Presence Of Noise, *Proceedings of the IRE* **37**, 10-21, 1949.

[268] A. Jessop, *Informed Assessments: An Introduction to Informtion, Entropy and Statistics*, New York : Ellis Horwood, 1995.

[269] J. P. W. Pluim, "Mutual Information Based registration of Medical Images," Utrecht, The Netherlands, Thesis Date 2000.

[270] S. Kullback and R. A. Leibler, On information and sufficiency, *Annals of Math Stats* **22**, 79-86, 1951.

[271] Y. M. Zhu, Volume image registration by cross-entropy optimization, *IEEE Trans Med Imaging* **21**, 174-180, 2002.

[272] D. Ros, M. Espinosa, J. F. Setoain, C. Falcon, F. J. Lomena, and J. Pavia, Evaluation of algorithms for the registration of 99Tcm-HMPAO brain SPET studies, *Nucl Med Commun* **20**, 227-236, 1999.

[273] B. A. Ardekani, M. Braun, I. Kanno, and B. F. Hutton, Automatic detection of intradural spaces in MR images, *J Comput Assist Tomogr* **18**, 963-969, 1994.

[274] P. Thevenaz and M. Unser, An efficient mutual information optimizer for multiresolution image registration, *International Conference on Image Processing ICIP 98*, pp.833-837, 1998.

[275] J. P. W. Pluim, J. B. Antoine Maintz, and M. A. Viegever, "Mutual Information Matching in Multiresolution Contexts," in *Image and Vision Computing*, vol. 19(1-2), F. Pernus, S. Kovacic, H. S. Stiehl, and M. A. Viergever, Eds., 2001, pp.45-52.

[276] D. Sarrut and S. Miguet, Fast 3D image transformations for registration procedures, *International Conference on Image Analysis and Processing Proceedings*, pp.446-451, 1999.

[277] J. L. Bernon, V. Boudousq, J. F. Rohmer, M. Fourcade, M. Zanca, M. Rossi, and D. Mariano-Goulart, A comparative study of Powell's and Downhill Simplex algorithms for a fast multimodal surface matching in brain imaging, *Computerized Medical Imaging and Graphics* **25**, 287-297, 2001.

[278] M. Y. Wang, C. R. Maurer, Jr., J. M. Fitzpatrick, and R. J. Maciunas, An automatic technique for finding and localizing externally attached markers in CT and MR volume images of the head, *IEEE Trans Biomed Eng* **43**, 627-637, 1996.

[279] Y. E. Erdi, B. W. Wessels, R. DeJager, A. K. Erdi, L. Der, Y. Cheek, R. Shiri, E. Yorke, R. Altemus, V. Varma, and et al., A new fiducial alignment system to overlay abdominal computed tomography or magnetic resonance anatomical images with radiolabeled antibody single-photon emission computed tomographic scans, *Cancer* **73**, 923-931, 1994.

[280] E. I. Parsai, K. M. Ayyangar, R. R. Dobelbower, and J. A. Siegel, Clinical fusion of three-dimensional images using Bremsstrahlung SPECT and CT, *J Nucl Med* **38**, 319-324, 1997.

[281] H. Inagaki, T. Kato, M. Tadokoro, K. Ito, H. Fukatsu, T. Ota, T. Isomura, M. Nishino, and T. Ishigaki, Interactive fusion of three-dimensional images of upper abdominal CT and FDG PET with no body surface markers, *Radiat Med* **17**, 155-163, 1999.

[282] R. J. Hamilton, M. J. Blend, C. A. Pelizzari, B. D. Milliken, and S. Vijayakumar, Using vascular structure for CT-SPECT registration in the pelvis, *J Nucl Med* **40**, 347-351, 1999.

[283] J. A. Patton, D. Delbeke, and M. P. Sandler, Image fusion using an integrated, dual-head coincidence camera with X- ray tube-based attenuation maps, *J Nucl Med* **41**, 1364-1368, 2000.

[284] M. S. Rosenthal, J. Cullom, W. Hawkins, S. C. Moore, B. M. W. Tsui, and M. Yester, Quantitative SPECT imaging: a review and recommendations by the focus committee of the Society of Nuclear Medicine Computer and Instrumentation Council, *J Nucl Med* **36**, 1489-1513, 1995.

[285] L. Bouwens, R. Van de Walle, J. Nuyuts, M. Koole, Y. D'Asseler, S. Vandenberghe, I. Lemahieu, and R. A. Dierckx, Image-correction techniques in SPECT, *Comput Med Imaging Graph* **25**, 117-126, 2001.

[286] H. Anger and D. Davis, Gamma-ray detection efficiency and image resolution in sodium iodide, *Rev Sci Instrum* **35**, 693, 1964.

[287] S. J. Glick, The effect of intrinsic spatial resolution on the quantitative accuracy of SPECT imaging, *IEEE Trans Nucl Sci* **46**, 1009-1015, 1999.

[288] M. Jahns, The influence of penetrating radiation on collimator performance, *Phys Med Biol* **26**, 113-124, 1981.

[289] J. Fleming and A. Alaamer, Influence of collimator characteristics on quantification in SPECT, *J Nucl Med* **37**, 1832-1836, 1996.

[290] G. Muehllehner and J. Karp, Advances in SPECT and PET, *IEEE Trans Nucl Sci* **35**, 639-643, 1988.

[291] E. D. Williams, Towards perfection in radionuclide imaging, *Phys Med Biol* **35**, 1147-1150, 1990.

[292] L. A. Shepp and Y. Vardi, Maximum likelihood reconstruction for emission tomography, *IEEE Trans Med Imag* **1**, 113-122, 1982.

[293] N. Natterer, *The mathematics of computerized tomography*, New York: Wiley, 1986.

[294] E. S. Chornoboy, C. J. Chen, M. I. Miller, T. R. Miller, and D. L. Snyder, An evaluation of maximum likelihood reconstruction for SPECT, *IEEE Trans Med Imag* **9**, 99-110, 1990.

[295] A. C. Kak and M. Slaney, *Principles of computerized tomographic imaging*, New York: IEEE Press, 1988.

[296] S. Vandenberghe, Y. D'Asseler, R. Van de Walle, T. Kauppinen, M. Koole, L. Bouwens, K. Van Laere, I. Lemahieu, and R. A. Dierckx, Iterative reconstruction algorithms in nuclear medicine, *Comp Med Imag Graph* **25**, 105-111, 2001.

[297] G. L. Zeng, Image reconstruction - a tutorial, *Comp Med Imag Graph* **25**, 97-103, 2001.

[298] P. P. Bruyant, Analytic and iterative reconstruction algorithms in SPECT, *J Nucl Med* **43**, 1343-1358, 2002.

[299] A. Seret, Median root prior and ordered subsets in bayesian image reconstruction of single-photon emission tomography, *Eur J Nucl Med* **25**, 215-219, 1998.

[300] G. T. Herman, *Image reconstruction from projections: the fundamentals of computerized tomography*, New York: Academic Press, 1980.

[301] K. Van Laere, M. Koole, I. Lemahieu, and R. Dierckx, Image filtering in single-photon emission computed tomography:pronciples and applications, *Comp Med Imag Graph* **25**, 127-133, 2001.

[302] D. L. Snyder, M. I. Miller, L. J. Thomas, and D. G. Politte, Noise and edges artifacts in maximum-likelihood reconstruction for emission tomography, *IEEE Trans Med Imag* **6**, 313-319, 1987.

[303] G. Germano, Technical aspects of myocardial SPECT imaging, *J Nucl Med* **42**, 1499-1507, 2001.

[304] V. Y. Panin, G. L. Zeng, and G. T. Gullberg, Total variation regulated EM algorithm, *IEEE Trans Nucl Sci* **46**, 2202-2210, 1999.

[305] F. J. Beekman, C. Kamphuis, M. A. King, P. P. van Rijk, and M. A. Viegever, Improvement of image resolution and quantitative accuracy in clinical single photon emission computed tomography, *Comp Med Imag Graph* **25**, 135-146, 2001.

[306] K. Lange and R. Carson, EM reconstruction algorithms for emission and transmission tomography, *J Comput Assist Tomogr* **8**, 306-316, 1984.

[307] Z. Liang and H. Hart, Bayesian reconstruction in emission computerized tomography, *IEEE Trans Nucl Sci* **35**, 788-792, 1988.

[308] P. J. Green, Bayesian reconstruction from emission tomography data using a modified EM algorithm, *IEEE Trans Med Imaging* **9**, 84-93, 1990.

[309] J. Llacer and E. Veklerov, Feasible images and practical stopping rules for iterative algorithms in emission tomography, *IEEE Trans Med Imaging* **6**, 37-51, 1989.

[310] E. T. P. Slijpen and F. J. Beekman, Comparison of post-filtering and filtering between iterations for SPECT reconstruction, *IEEE Trans Nucl Sci* **46**, 2233-2238, 1999.

[311] B. F. Hutton and Y. H. Lau, Application of distance-dependent resolution compensation and post-reconstruction filtering for myocardial SPECT, *Phys Med Biol* **43**, 1679-1693, 1998.

[312] C. Kamphuis, F. J. Beekman, and B. F. Hutton, Optimal collimator hole dimensions for half cone-beam brain SPECT, *International Meeting on Fully three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine Proceedings*, pp.271-275, 1999.

[313] Y. H. Lau, B. F. Hutton, and F. J. Beekman, Choice of collimator for cardiac SPET when resolution compensation is included in iterative reconstruction, *Eur J Nucl Med* **28**, 39-47, 2001.

[314] L. T. Chang, A method for attenuation correction in radionuclide computed tomography, *IEEE Trans Nucl Sci* **25**, 638-643, 1978.

[315] K. Murase, H. Itoh, H. Mogami, M. Kawamura, A. Iio, and K. Hamamoto, A comparative study of attenuation correction algorithms in single photon emission computed tomography (SPECT), *Eur J Nucl Med* **13**, 55-62, 1987.

[316] D. L. Bailey, B. F. Hutton, and P. J. Walter, Improved SPECT using simultaneous emission and transmission tomography, *J Nucl Med* **28**, 844-851, 1987.

[317] P. Tan, D. L. Bailey, S. R. Meikle, S. Eberl, and R. R. Fulton, A scanning line source for simultaneous emission and transmission measurements in SPECT, *J Nucl Med* **34**, 1752-1760, 1993.

[318] S. R. Meikle, B. F. Hutton, and D. L. Bailey, A transmission-dependent method for scatter correction in SPECT, *J Nucl Med* **35**, 360-367, 1994.

[319] F. J. Beekman and M. A. Viergever, Fast SPECT simulation including object shape dependent scatter, *IEEE Trans Med Imag* **14**, 271-282, 1995.

[320] B. F. Hutton, A. Osiecki, and S. R. Meikle, Transmission-based scatter correction of 180 degrees myocardial single-photon emission tomographic studies, *J Nucl Med* **23**, 1300-1308, 1996.

[321] F. J. Beekman, C. Kamphuis, and M. A. Viegever, Improved SPECT quantitation using fully three-dimensional iterative spatially variant scatter response compensation, *IEEE Trans Med Imaging* **15**, 491-499, 1996.

[322] G. L. Zeng, C. Bai, and G. T. Gullberg, A projector/backprojector with slice-to-slice blurring for efficient three-dimensional scatter modeling, *IEEE Trans Med Imag* **18**, 722-732, 1999.

[323] G. El Fakhri, S. C. Moore, and P. Maksud, A new scatter compensation method for Ga-67 imaging using artificial neural networks, *IEEE Trans Nucl Sci* **48**, 799-804, 2001.

[324] G.-Z. Yang, P. Burger, J. Panting, P. D. Gatehouse, D. Rueckert, D. J. Pennell, and D. N. Firmin, Motion and deformation tracking for short-axis echo-planar myocardial perfusion imaging, *Med Image Anal* **2**, 285-302, 1998.

[325] R. R. Fulton, S. Eberl, S. R. Meikle, B. F. Hutton, and M. Braun, A practical 3D tomographic method for correcting patient head motion in clinical SPECT, *IEEE Trans Nucl Sci* **46**, 667-672, 1999.

[326] A. J. Da Silva, H. R. Tang, M. C. Wu, and B. H. Hasegawa, Absolute quantitation of myocardial activity in phantoms, *IEEE Trans Nucl Sci* **46**, 659-666, 1999.

[327] A. J. Da Silva, H. R. Tang, K. H. Wong, M. C. Wu, M. W. Dae, and B. H. Hasegawa, Absolute in vivo quantitation of myocardial activity, *IEEE Trans Nucl Sci* **47**, 1093-1098, 2000.

[328] M. Koole, R. Van de Walle, K. Van Laere, Y. D'Asseler, S. Vandenberghe, I. Lemahieu, and R. A. Dierckx, Study of the quantification of FBP SPECT images with a correction for partial volume effects, *IEEE Trans Nucl Sci* **49**, 69-73, 2002.

[329] K. J. Donohoe, R. E. Henkin, H. D. Royal, M. L. Brown, B. D. Collier, R. E. O'Mara, and R. F. Carretta, Procedure guideline for bone scintigraphy:1.0, *J Nucl Med* **37**, 1903-1906, 1996.

[330] J. E. Seabold, C. J. Palestro, M. L. Brown, F. L. Datz, L. A. Forstrom, and B. S. Greenspan, Procedure guideline for gallium scintigraphy in inflammation, *J Nucl Med* **38**, 994-997, 1997.

[331] S. P. Bartold, K. J. Donohoe, J. W. Fletcher, T. P. Haynie, R. E. Henkin, E. B. Silberstein, H. D. Royal, and A. Van den Abbeele, Procedure guideline for gallium scintigraphy in the evaluation of malignant disease, *J Nucl Med* **38**, 990-994, 1997.

[332] S. C. Moore, S. P. Kijewski, S. P. Muller, and B. L. Holman, SPECT image noise power: effects of nonstationary projection noise and attenuation compensation, *J Nucl Med* **29**, 1704-1709, 1988.

[333] J. S. Liow and S. C. Strother, Practical tradeoffs between noise, quantitation, and number of iterations for maximum likelihood-based reconstruction, *IEEE Trans Med Imag* **10**, 563-572, 1991.

[334] H. H. Barrett, D. W. Wilson, and B. M. W. Tsui, Noise properties of EM algorithm,I: theory, *Phys Med Biol* **39**, 833-846, 1994.

[335] D. W. Wilson, H. H. Barrett, and B. M. W. Tsui, Noise properties of EM algorithm,II: Monte Carlo simulations, *Phys Med Biol* **39**, 845-871, 1994.

[336] E. J. Soares, S. J. Glick, and M. A. King, Noise characterization of combined Bellini-type attenuation correction and frequency-distance principle restoration filtering [SPECT], *IEEE Trans Nucl Sci* **43**, 3278-3290, 1996.

[337] P. P. Bruyant, J. Sau, and J.-J. Mallet, Streak artifact reduction in filtered backprojection using a level line-based interpolation method, *J Nucl Med* **41**, 1913-1919, 2000.

[338] E. J. Soares, C. L. Byrne, and S. J. Glick, Noise characterization of block-iterative reconstruction algorithms. I. Theory, *IEEE Trans Med Imag* **19**, 261-270, 2000.

[339] D. W. Wilson and B. M. W. Tsui, Noise properties of filtered-backprojection and ML-EM reconstructed emission tomographic images, *IEEE Trans Nucl Sci* **40**, 1198-1203, 1993.

[340] D. Blocklet, A. Seret, N. Popa, and A. Schoutens, Maximum-likelihood reconstruction with ordered subsets in bone SPECT, *J Nucl Med* **40**, 1978-1984, 1999.

[341] R. G. Wells, P. H. Simkin, P. F. Judy, M. A. King, P. H. Pretorius, H. C. Gifford, and P. Schneider, Maximizing the detection and localization of Ga-67 tumors in thoracic SPECT MLEM(OSEM) reconstructions, *IEEE Trans Nucl Sci* **46**, 1191-1198, 1999.

[342] R. G. Wells, M. A. King, P. H. Simkin, P. F. Judy, A. B. Brill, H. C. Gifford, R. Licho, P. H. Pretorius, P. B. Schneider, and D. W. Seldin, Comparing filtered backprojection and ordered-subsets expectation maximization for small-lesion detcetion and localization in ^{67}Ga SPECT, *J Nucl Med* **41**, 1391-1399, 2000.

[343] B. Nelson, R. L. Hayes, and et.al., Distribution of Gallium in human tissue after intravenous administration, *J Nucl Med* **13**, 92-100, 1972.

[344] S. L. Bacharach and I. Buvat, Attenuation correction in cardiac PET and SPECT, *J Nucl Cardiol* **2**, 246-255, 1995.

[345] P. H. Pretorius, M. A. King, T.-S. Pan, and B. F. Hutton, Attenuation correction strategies for multi-energy photon emitters using SPECT, *IEEE Trans Nucl Sci* **44**, 1323-1328, 1997.

[346] J. A. Little and D. J. Hawkes, The registration of multiple medical images acquired from a single subject: why, how, what next?, *Stat Methods Med Res* **6**, 239-65., 1997.

[347] K. Suga, N. Matsunaga, Y. Kawakami, and M. Furukawa, Phantom study of fusion image of CT and SPECT with body-contour generated from external Compton scatter sources, *Ann Nucl Med* **14**, 271-277, 2000.

[348] B. F. Hutton and Y. H. Lau, How critical is subset balance for OSEM reconstruction in myocardial SPECT?, *J Nucl Med* **57p**, 385, 1997.

[349] R. L. Wahl, L. E. Quint, R. D. Cieslak, A. M. Aisen, R. A. Koeppe, and C. R. Meyer, "Anatometabolic" tumor imaging: fusion of FDG PET with CT or MRI to localize foci of increased activity, *J Nucl Med* **34**, 1190-1197, 1993.

[350] D. L. Wilson, A. Carrillo, L. Zheng, A. Genc, J. L. Duerk, and J. S. Lewin, Evaluation of 3D image registration as applied to MR-guided thermal treatment of liver cancer, *J Magn Reson Imaging* **8**, 77-84, 1998.

[351] J. Skalski, R. L. Wahl, and C. R. Meyer, Comparison of mutual information-based warping accuracy for fusing body CT and PET by 2 methods: CT mapped onto PET emission scan versus CT mapped onto PET transmission scan, *J Nucl Med* **43**, 1184-1187, 2002.

[352] D. K. Chung, Y. H. Lau, T. Kitosos, B. F. Hutton, and R. B. Howman-Giles, Application of multi-modality imaging coregistration in a paediatric nuclear medicine unit, *Nucl Med Commun* **22**, 920-921, 2001.

[353] B. C. Vemuri, J. Ye, Y. Chen, and C. M. Leonard, Image registration via level-set motion: Applications to atlas-based segmentation, *Med Image Anal* **7**, 1-20, 2003.

[354] S. L. Hartmann, M. H. Parks, P. R. Martin, and B. M. Dawant, Automatic 3-D segmentation of internal structures of the head in MR images using a combination of similarity and free-form transformations. II. Validation on severely atrophied brains, *IEEE Trans Med Imag* **18**, 917-926, 1999.

[355] M. Ferrant, A. Nabavi, B. Macq, P. M. Black, F. A. Jolesz, R. Kikinis, and S. K. Warfield, Serial registration of intraoperative MR images of the brain, *Med Image Anal* **6**, 337-359, 2002.

[356] G. Berks, A. Ghassemi, and D. G. von Keyserlingk, Spatial registration of digital brain atlases based on fuzzy set theory, *Comput Med Imaging Graph* **25**, 1-10, 2001.

[357] M. Xu and W. L. Nowinski, Talairach-Tournoux brain atlas registration using a metalforming principle-based finite element method, *Med Image Anal* **5**, 271-279, 2001.

[358] D. Shen, Z. Lao, J. Zeng, W. Zhang, I. A. Sesterhenn, L. Sun, J. W. Moul, E. H. Herskovits, G. Fichtinger, and C. Davatzikos, Optimized prostate biopsy via a statistical atlas of cancer spatial distribution, *Medical Image Analysis* **8**, 139-150, 2004.

[359] D. Rey, G. Subsol, H. Delingette, and N. Ayache, Automatic detection and segmentation of evolving processes in 3D medical images: Application to multiple sclerosis, *Med Image Anal* **6**, 163-179, 2002.

[360] G. Subsol, J.-P. Thirion, and N. Ayache, A scheme for automatically building three-dimensional morphometric anatomical atlases: application to a skull atlas, *Med Image Anal* **2**, 37-60, 1998.

[361] S. Warfield, A. Robatino, J. Dengler, F. Jolesz, and R. Kikinis, "Nonlinear registration and template driven segmentation," in *Brain Warping (Ch 4)*, vol., A. W. Toga, Eds., New York: Academic Press, 1999.

[362] R. Szeliski, Bayesian modeling of uncertainty in low-level vision, *Int J Comput Vision* **5**, 271-301, 1990.

[363] J. C. Gee, C. Barillot, L. L. Briquer, D. R. Haynor, and R. Bajcsy, Matching structural images of the human brain using statistical and geometrical images features, *Proc. SPIE 2359:Visualization in Biomedical Computing*, pp.191-201, 1994.

[364] J. C. Gee, L. L. Briquer, C. Barillot, and D. R. Haynor, "Probabilistic matching of images," in *Proc. 14th Int. Conf. Information Processing in Medical Imaging: Computational Imaging and Vision 3*, vol., Bizais, Barillot, and Y. C. R. Di Paols, Eds., Boston: Kluwer, 1995, pp.113-125.

[365] R. Szeliski and S. Lavallee, Matching 3-D anatomical surface with non-rigid deformations using octree-splines, *Int J Comput Vision* **18**, 171-186, 1996.

[366] J. Ruiz-Alzola, C. F. Westin, S. K. Warfield, C. Alberola, S. Maier, and R. Kikinis, Nonrigid registration of 3D tensor medical data, *Med Image Anal* **6**, 143-161, 2002.

[367] J.-P. Thirion, Image matching as a diffusion process: an analogy with Maxwell's demons, *Med Image Anal* **2**, 243-260, 1998.

[368] G. E. Christensen, S. C. Joshi, and M. I. Miller, Volumetric transformation of brain anatomy, *IEEE Trans Med Imaging* **16**, 864-877, 1997.

[369] G. Wollny and F. Kruggel, Computional cost of nonrigid registartion algorithms based on fluid dynamics, *IEEE Trans Med Imaging* **21**, 946-952, 2002.

[370] A. Roche, X. Pennec, G. Malandain, and N. Ayache, Rigid registration of 3-D ultrasound with MR images: a new approach combining intensity and gradient information, *IEEE Trans Med Imaging* **20**, 1038-1049, 2001.

[371] Y. H. Lau, M. Braun, and B. F. Hutton, "Non-rigid 3D image registration using regionally constrained matching and the correlation ratio," in *Proc. International Workshop on Biomedical Image Registration*, vol., F. Pernus, S. Kovacic, H. S. Stiehl, and M. A. Viergever, Eds., Bled, Slovenia, 1999, pp.137-148.

[372] Y.-C. Tai, K. P. Lin, C. K. Hoh, S. C. H. Huang, and E. J. Hoffman, Utilization of 3-D elastic transformation in the registration of chest X-ray CT and whole body PET, *IEEE Trans Nucl Sci* **44**, 1606-1612, 1997.

[373] D. Terzopoulos, Image analysis using multigrid relaxation methods, *IEEE Trans Patt Anal Mach Intell* **2**, 129-139, 1986.

[374] P. Thevenaz and M. Unser, A pyramid approach to sub-pixel image fusion based on mutual information, *International Conference on Image Processing Proceedings*, pp.265-268, 1996.

[375] L. Thurfjell, J. Andersson, M. Pagani, C. Jonsson, R. Hult, R. Lundqvist, A. Wagner, H. Jacobsson, and S. Larsson, Automatic detection of hypoperfused areas in SPECT brain scans, *IEEE Trans Nucl Sci* **45**, 2149-2154, 1998.

[376] N. Hata, T. Dohi, S. Warfield, I. W. Wells, R. Kikins, and F. A. Jolesz, "Multimodality deformable registration of pre- and intraoperative images for MRI-guided Brain Surgery," in *Proceedings of the First International Conference on Medical Image Computing and Computer-Assisted Intervention MICCAI'98, Lecture Notes in Computer Science 1496*, vol., W. M. Wells, A. Colchester, and S. Delp, Eds., Springer-Verlag, 1998, pp.1067-1074.

[377] P. J. Kostelec, J. B. Weaver, and D. M. Healy, Jr., Multiresolution elastic image registration, *Med Phys* **25**, 1593-1604, 1998.

[378] P. Cachier, X. Pennec, and N. Ayache, Fast non rigid matching by gradient descent: study and improvements of the "demons" algorithm, Inst. National de Recherche d'Informatique et d'Auto (INRIA), France, Report No. 3706, 1999.

[379] A. Guimond, A. Roche, N. Ayache, and J. Meunier, Multimodal brain warping using the demons algorithm and adaptative intensity corrections, Inst. National de Recherche d'Informatique et d'Auto (INRIA), France, Report No. 3796, 1999.

[380] B. Kim, J. L. Boes, K. A. Frey, and C. R. Meyer, Mutual information for automated unwarping of rat brain autoradiographs, *Neuroimage* **5**, 31-40, 1997.

[381] C. R. Meyer, J. L. Boes, B. Kim, P. H. Bland, G. L. Lecarpentier, J. B. Fowlkes, M. A. Roubidoux, and P. L. Carson, Semiautomatic registration of volumetric ultrasound scans, *Ultrasound in Medicine & Biology* **25**, 339-347, 1999.

[382] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank, PET-CT image registration in the chest using free-form deformations, *IEEE Trans Med Imag* **22**, 120-128, 2003.

[383] J. S. Duncan and N. Ayache, Medical image analysis: progress over two decades and the challenges ahead, *IEEE Trans Patt Anal Mach Intell* **22**, 85-106, 2000.

[384] M. E. Leventon, W. Eric, and L. Grimson, Multi-modal volume registration using joint intensity distributions, Proceedings of the First International Conference on Medical Image Computing and Computer-Assisted Intervention MICCAI'98, Lecture Notes in Computer Science 1496, pp.1057-1066, 1998.

[385] D. W. Townsend and T. Beyer, A combined PET/CT scanner: the path to true image fusion, *Br J Radiol* **75 Spec No**, S24-30, 2002.

[386] R. Bar-Shalom, N. Yefremov, L. Guralnik, D. Gaitini, A. Frenkel, A. Kuten, H. Altman, Z. Keidar, and O. Israel, Clinical performance of PET/CT in evaluation of cancer: additional value for diagnostic imaging and patient management, *J Nucl Med* **44**, 1200-1209, 2003.

[387] D. W. Townsend, J. P. J. Carney, J. T. Yap, and N. C. Hall, PET/CT today and tomorrow, *J Nucl Med* **45**, 4S-14S, 2004.

[388] T. Nishioka, T. Shiga, H. Shirato, E. Tsukamoto, K. Tsuchiya, T. Kato, K. Ohmori, A. Yamazaki, H. Aoyama, S. Hashimoto, T. C. Chang, and K. Miyasaka, Image fusion between 18FDG-PET and MRI/CT for radiotherapy planning of oropharyngeal and nasopharyngeal carcinomas, *Int J Radiat Oncol Biol Phys* **53**, 1051-1057, 2002.

[389] O. Sipila, P. Nikkinen, S. Savolainen, M. L. Granstrom, E. Gaily, V. P. Poutanen, H. Pohjonen, and K. Liewendahl, Transmission imaging for registration of ictal and interictal single- photon emission tomography, magnetic resonance imaging and electroencephalography, *Eur J Nucl Med* **27**, 202-205, 2000.

[390] T. Makela, Q. C. Pham, P. Clarysse, J. Nenonen, J. Lotjonen, O. Sipila, H. Hanninen, K. Lauerma, J. Knuuti, T. Katila, and I. E. Magnin, A 3-D model-based registration approach for the PET, MR and MCG cardiac data fusion, *Medical Image Analysis* **7**, 377-389, 2003.

Publications list

Articles in refereed journals

1. Brian F Hutton, Michael Braun, Lennart Thurfjell, Dennys YH Lau, Image registration: an essential tool for nuclear medicine, *Eur J Nucl Med,* vol.29, pp.559-577, 2002

   My contribution to the published paper focused on reviewing journal papers related to image registration and preparing diagrams.

2. YH Lau, M Braun and BF Hutton, Non-Rigid image registration using a median-filtered coarse-to-fine displacement field and a symmetric correlation ratio, *Phys Med Biol,* vol.46, pp.1297-1319, 2001.

   My contribution to the published paper focused on developing, implementing, validating the proposed algorithm, and writing up.

3. L Thurfjell, YH Lau and BF Hutton, Improved efficiency of images registration by using mutual information, *Eur J Nucl Med,* vol.27, pp.847-856, 2000.

   My contribution to the published paper focused on developing of a strategy for improving registration efficiency, implementing and testing the strategy.

4. LR Barnden, R Kwiatek, YH Lau, BF Hutton, L Thurfjell and CC Rowe, Validation of fully automatic brain SPECT to MR co-registration, *Eur J Nucl Med,* vol.27, pp.147-154, 2000.

   My contribution to the published paper focused on validating the proposed algorithm with the supplied image data set.

Chapter 2 mentions the work related to the review of medical image registration in (1). Chapter 7 refers to the work in (2) on the development of a non-rigid registration method. Chapter 8 extends the validation of 3D intra-modality non-rigid registration in (2) to inter-modality image. The work of Thurfjell *et al.* in (3) is not specifically included but the results are broadly implemented. In Chapter 4 the work of Barnden *et al.* in (4) is mentioned and was essentially repeated with extensions, which includes a comparison of various sampling strategies between different similarity measures.

Articles in proceedings of international conferences

- YH Lau, M Braun and BF Hutton, Non-rigid 3D image registration using regionally constrained matching and correlation ratio, in F.Pernus, S.Kovacic, H.S.Stiehl and M.A.Viergever (Eds.), Proc. International Workshop on Biomedical Image Registration 1999, Bled, Slovenia. pp.137-148, 1999.
- YH Lau, BF Hutton, M Braun, An approach to non-rigid registration, Workshop for collaborative Research on Emerging computational methods for application of SPECT in Cardiology, Neurology and Oncology, University Hospital, Utrecht, Netherlands, 7-10 September, 1998.

Conferences abstracts

- R. Howman-Giles, D. Chung, Y.H. Lau, T. Kitsos, R. Uren, and B. Hutton, Does mulit-modality image co-registration impact on clinical management in paediatric diseases? *J Nucl Med*, vol.43(5S), pp.84-85, 2002.

- D.K. Chung, Y.H. Lau, T. Kitsos, B.F. Hutton, and R.B. Howman-Giles, Application of multi-modality imaging coregistration in a paediatric nuclear medicine unit. *Nucl Med Commun*, vol.22(8), pp.920-921, 2001.

- D.K. Chung, Y.H. Lau, T. Kitsos, B.F. Hutton and R.B. Howman-Giles, Application of multi-modality image coregistration in a paediatric nuclear medicine unit, 31st Annual Scientific Meeting of the Australian and New Zealand Society of Nuclear Medicine - Hobart, Tasmania, Australia, 18-22 May 2001.

- YH Lau, M Braun and BF Hutton, Registration of SPECT and CT abdominal images using a symmetric correlation ratio. 30th Annual Scientific Meeting, The Australian and New Zealand Society of Nuclear Medicine, Adelaide, 2000.

- L Thurfjell, YH Lau and BF Hutton, Multi-modality registration of brain scans using mutual information. Proceedings of 29th Annual Scientific Meeting, The Australian and New Zealand Society of Nuclear Medicine, Perth. *Nucl Med Commun*, vol.20(4), pp.373 (abs), 1999.

- YH Lau, M Braun and BF Hutton, Deformable grid model for non-rigid registration. Proceedings of 30th Annual Scientific Meeting, The Australian and New Zealand Society of Nuclear Medicine, Perth. *Nucl Med Commun*, vol.20(4), pp.373 (abs), 1999.

- YH Lau, M Braun and BF Hutton, Non-rigid 3D image registration using regionally constrained matching and correlation ratio. Sixth National Workshop Experimental Radiation Oncology, St. George Hospital, Sydney, Australia, 1999.

Publications not related to this thesis

Articles in refereed journals

- BF Hutton, AZ Kyme, YH Lau, DW Skerrett and RR Fulton, A hybrid 3-D reconstruction/registration algorithm for correction of head motion in emission tomography, *IEEE Trans Nucl Sci,* vol. 49(1), pp.188-194, 2002.

- YH Lau, BF Hutton, FJ Beekman, Choice of collimator for cardiac SPECT when resolution compensation is included in iterative reconstruction, *Eur J Nucl Med* vol.28, pp.39-47, 2000.

- BF Hutton and YH Lau, Application of distance-dependent resolution compensation and post-reconstruction filtering for myocardial SPECT, *Phys Med Biol*, vol.43, pp.1679-1693, 1998.

Articles in proceedings of international conferences

- YH Lau, BF Hutton and FJ Beekman, Choice of collimator for cardiac SPECT when 3D resolution compensation is included in OSEM reconstruction. Proc. International meeting on fully three-dimensional image reconstruction in radiology and nuclear medicine, 1999. Utrecht, Netherlands, F.Beekman, M.Defrise, and M.Viergever (Eds.), pp.347-350. 1999.

- YH Lau, BF Hutton, Resolution compensation and post-reconstruction filtering for myocardial SPECT, Workshop for collaborative Research on Emerging computational methods for application of SPECT in Cardiology, Neurology and Oncology, University Hospital, Utrecht, Netherlands, 7-10 September, 1998.

- YH Lau, BF Hutton, Choice of collimator for OSEM reconstruction with resolution compensation, Workshop for collaborative Research on Emerging computational methods for application of SPECT in Cardiology, Neurology and Oncology, University Hospital, Utrecht, Netherlands, 7-10 September, 1998.

- BF Hutton, YH Lau, Further improvement in the efficiency of model based scatter estimation for use in OSEM SPECT reconstruction, Proceedings of 44th Annual Meeting of the Society of Nuclear Medicine, San Antonio, Texas, *J Nucl Med*, vol.38 (5), pp.88, 1997.

- BF Hutton, YH Lau, How critical is subset balance for OSEM reconstruction in myocardial SPECT? Proceedings of 44th Annual Meeting of the Society of Nuclear Medicine, San Antonio, Texas, *J Nucl Med*, vol.38 (5), pp.57, 1997

- BF Hutton, YH Lau, Myocardial SPECT reconstruction using EM and OSEM with compensation for distance dependent resolution and post-reconstruction filtering. Proceedings of 44th Annual Meeting of the Society of Nuclear Medicine, San Antonio, Texas, *J Nucl Med*, vol.38 (5), pp.20, 1997

Conferences abstracts

- RL Hatton, BF Hutton, S Som and YH Lau, Technical issues relating to the statistical parametric mapping of brain SPECT studies. 30th Annual Scientific Meeting, The Australian and New Zealand Society of Nuclear Medicine, Adelaide, 2000.

- YH Lau, BF Hutton, S Evans, Choice of collimator for cardiac SPECT when compensation for resolution is included in OSEM reconstruction: validation using a thorax phantom. 30th Annual Scientific Meeting, The Australian and New Zealand Society of Nuclear Medicine, Adelaide, 2000.

- YH Lau, BF Hutton, The case for use of a general purpose collimator for cardiac SPECT. Proceedings of 28th Annual Scientific Meeting, Australia and New Zealand Society of Nuclear Medicine, Melbourne, April 4-8, pp.59, 1998.

Appendix

Anatomical labels of the Zubal phantom are shown below:

```
label(0)='outside_phantom'
label(1)='skin'
label(2)='brain'
label(3)='spinal_cord'
label(4)='skull'
label(5)='spine'
label(6)='dens of axis'
label(7)='jaw bone'
label(8)='skeletal_muscle'
label(9)= 'lacrimal glands'
label(10)= 'spinal canal'
label(11)= 'hard palate'
label(12)= 'cerebellum'
label(13)= 'tongue'
label(14)= 'pharynx'
label(15)= 'esophagus'
label(16)= 'medulla oblongota'
label(17)= 'fat'
label(18)= 'blood pool'
label(19)= 'bone marrow'
label(20)= 'pons'
label(21)= 'trachea'
label(22)= 'cartilage'
label(23)= 'uncus(ear bones)'
label(24)= 'sinuses/mouth cavity'
label(25)= 'optic nerve'
label(26)= 'cerebral falx'
label(27)= 'eye'
label(28)= 'lens'
label(29)= 'cerebral aquaduct'
label(30)= 'teeth'
label(31)='rib cage & sternum'
label(32)='pelvis'
label(33)='long bones'
label(34)='lungs'
label(35)='heart'
label(36)='liver'
label(37)='gall bladder'
label(38)='kidney'
label(39)='stomach'
label(40)='small bowel'
label(41)='colon'
label(42)='pancreas'
label(43)='adrenals'
label(44)='gas (bowel)'
label(45)='fluid (bowel)'
label(46)='lymph nodes'
label(47)='thyroid'
label(48)='spleen'
label(49)='urine'
label(50)='feces'
label(51)='testes'
label(52)='prostate'
label(53)='rectum'
label(54)='diaphragm'
label(55)='bladder'
label(56)='lesion'
```