UNIVERSITY OF TECHNOLOGY, SYDNEY

Surface-type Classification in Structured Planar Environments under Various Illumination and Imaging Conditions

by

Andrew Wing Keung To

A thesis submitted in partial fulfillment for the
degree of Doctor of Philosophy

in the
Faculty of Engineering and IT
Electrical, Mechanical and Mechatronic Systems Group
Centre for Autonomous Systems

July 2015

Declaration of Authorship

I, Andrew Wing Keung To , declare that this thesis titled, 'Surface-type Classification in Structured Planar Environments under Various Illumination and Imaging Conditions' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: ________________________________

Date: ________________________________

UNIVERSITY OF TECHNOLOGY, SYDNEY

Abstract

Faculty of Engineering and IT
Electrical, Mechanical and Mechatronic Systems Group

Doctor of Philosophy

by Andrew Wing Keung To

The recent advancement in sensing, computing and artificial intelligence, has led to the application of robots outside of the manufacturing factory and into field environments. In order for a field robot to operate intelligently and autonomously, the robot needs to build an environmental awareness, such as by classifying the different surface-types on a steel bridge structure. However, it is challenging to classify surface-types from images that are captured in a structurally complex environment under various illumination and imaging conditions. This is because colour and texture features extracted from these images can be inconsistent.

This thesis presents a surface-type classification approach to classify surface-types in a structurally complex three-dimensional (3D) environment under various illumination and imaging conditions. The approach proposes RGB-D sensing to provide each pixel in an image with additional depth information that is used by two developed algorithms. The first algorithm uses the RGB-D information along with a modified reflectance model to extract colour features for colour-based classification of surface-types. The second algorithm uses the depth information to calculate a probability map for the pixels being a specific surface-type. The probability map can identify the image regions that have a high probability of being accurately classified by a texture-based classifier.

A 3D grid-based map is generated to combine the results produced by colour-based classification and texture-based classification. It is suggested that a robot manipulator is used to position an RGB-D sensor package in the complex environments to capture the RGB-D images. In this way, the 3D position of each pixel is precisely known in a common global frame (robot base coordinate frame) and can be combined using a grid-based map to build up a rich awareness of the surrounding complex environment.

A case study is conducted in a laboratory environment using a six degree-of-freedom robot manipulator equipped with a RGB-D sensor package mounted to the end effector. The results show that the proposed surface-type classification approach provides an improved solution for vision-based classification of surface-types in a complex structural environment with various illumination and imaging conditions.

Acknowledgements

I would like to thank my supervisors Prof. Dikai Liu, and Dr Gavin Paul for their continual support and assistance throughout the course of my research. Your guidance and countless hours spent towards improving my research work has led to a more complete, quality thesis.

Thanks to the rest of the team at the Centre of Autonomous Systems, and Prof. Gamini Dissanayake, the head of the research centre, for providing an excellent environment that has facilitated great research interactions and exchange of ideas. Fellow research student Gibson Hu for providing encouragement and support throughout the course of the candidature.

I would finally like to thank my immediate family members Nelson To, Anita Luk, Anson To, grandparents and Ayesha Tang for believing in me to strive to do my very best.

This work is supported in part by the ARC Linkage Project: A robotic system for steel bridge maintenance, the Centre for Autonomous Systems (CAS), the NSW Roads and Maritime Services (RMS) and the University of Technology, Sydney (UTS).

Contents

List of Figures

List of Tables

Abbreviations

DOF	Depth of Field
FOV	Field of View
GLCM	Grey-Level Co-occurrence Matrices
IR	Infrared
PCA	Principal Component Analysis
RGB-D	Red, Green and Blue colour-space image with corresponding Depth image
SVM	Support Vector Machines
UTS	University of Technology, Sydney

Nomenclature

	General Formatting Style
$f(\cdots)$	A scalar valued function
$\mathbf{f}(\cdots)$	A vector valued function
$[\cdots]^T$	Transpose
$\lvert \cdot \rvert$	Absolute value
$\lVert \cdot \rVert$	Vector length and normalised vector
C	Covariance matrix
d	distance between two points
D	A diagonalised matrix
i	Index value in a list
n	Variable signifying the last index of a set or to refer to a count
P	Probability
(u, v)	Index values in a 2D array or an image
$\vec{v}$	A vector
Ω	2D image
τ	Threshold
θ	Angle between two directional vectors
	Specific Symbol Usage
${}^{o}T_e(\vec{Q})$	Homogenous transformation between the robot base coordinate frame and the end-effector at pose $\vec{Q}$
${}^{o}T_s$	Homogenous transformation matrix between the robot base coordinate frame and the sensor coordinate frame

${}^{e}T_s$	Homogenous transformation matrix between the robot end-effector coordinate frame and the sensor coordinate frame
${}^{s}T_c$	Homogenous transformation matrix between the sensor coordinate frame and the camera coordinate frame
C	Principle point of a pinhole camera model
d_c	Surface point-to-RGB camera coordinate origin distance
d_l	Light source-to-surface point distance
d_p	Plane of focus to camera distance
d_t	The viewing distance used to capture the training image dataset
F	Focal length of a camera
I_l	Light source intensity value
I_r	Reflected light source intensity value
K_d	Set notation for the diffused reflectance values
K_s	Set notation for the specular reflectance values
$\vec{l}$	The position vector of the light source relative to the RGB camera coordinate frame
n_t	Number of surface-types
P	A vector (or set) of 3D points or vertices
q	Robot manipulator's joint angle
$\vec{Q}$	Robot manipulator's joint angle vector, $[q_1, q_2, ...q_n]^T$
α	Bisector angle between $\vec{v_c}$ and $\vec{v_l}$
β_g	Gaussian blur coefficient
β_k	Skewing coefficient
β_s	Scaling coefficient
δ	Lens f-number
μ	Length of the voxel cube in the surface-type map
ω_v	Weighting factor applied to a voxel containing texture-based classification results
Ω_s	Depth image from the IR camera
Ω_c	Greyscale calibration image used to calculated the light source position

Ω_{cs}	An image of the greyscale calibration image Ω_c containing the specular reflectance region
Ω_{cd}	An image of the greyscale calibration image Ω_c containing the diffused reflectance region
Ω_t	A simulated texture pattern image
φ	Circle of confusion
σ	Surface roughness albedo
θ_c	Angle of incidence between the normal of a 3D surface point and the straight line between the surface point and the RGB camera coordinate origin
θ_l	Angle of incidence between the normal of a 3D surface point and the straight line between the surface point and the light source
θ_t	The viewing angle used to capture the training image dataset
τ_s	Pixel intensity threshold for identifying the specular reflectance region in an image
τ_d	Pixel intensity threshold for identifying the diffused reflectance region in an image
$\vec{v_\eta}$	Normal vector of a 3D surface point
$\vec{v_c}$	Direction vector between the surface point and the RGB camera coordinate origin
$\vec{v_l}$	Direction vector between the light source point and the RGB camera coordinate origin
$P(M_k)$	Discrete probability distribution of the surface-types for $k \in \{1, \ldots n_t\}$, given n_t number of surface-types.
$P(M_k\|E)$	Probability of surface-type state given the evidence E
$P(E\|M_k)$	Probability of an evidence given the surface-type
$P(E)$	Probability of evidence
P_{d_c}	Probability value of a pixel being a surface-type based on viewing distance
P_{θ_c}	Probability value of a pixel being a surface-type based on viewing angle

P_{d_c,θ_c}	Probability value of a pixel being a surface-type based on viewing distance and viewing angle

Combinations of Variables

(a_2, a_1, a_0)	Polynomial coefficients for camera radiometric response in the reflectance model
$\{D_{n_1}, D_{f_1}\}$	Depth of field threshold range
$\{D_{n_2}, D_{f_2}\}$	Spatial resolution threshold range
$(K_{d,R}, K_{d,G}, K_{d,B})$	Diffused reflectance value for each RGB colour channel
$(K_{s,R}, K_{s,G}, K_{s,B})$	Specular reflectance value for each RGB colour channel
(x_c, y_c, z_c)	Axes of RGB camera's 3D Cartesian coordinate frame
(x_o, y_o, z_o)	Axes of Robot base's 3D Cartesian coordinate frame
(x_e, y_e, z_e)	Axes of End-effector's 3D Cartesian coordinate frame
(x_s, y_s, z_s)	Axes of Depth sensor's 3D Cartesian coordinate frame
$(\tau_n, \tau_f, \tau_\theta)$	Threshold parameters to calculate an image pixel's probability of being a surface-type
(ω_1, ω_2)	Weighting coefficients to calculate an image pixel's probability of being a surface-type

Glossary of Terms

Complex environment	A 3D workspace that has multiple planar surfaces arranged in various positions and orientations.
Confusion matrix	A specific table that allows the visualisation of classification results. Each column of the matrix represents the instances in a predicted class, while each row represents the instances in an actual class.
Environmental awareness	In the context for a robot this can include but is not limited to the knowledge of, a geometric map of the environment that describes the location of surfaces and obstacles, and a semantic map that provides a label for objects, surface-types and locations within the environment.
Grid	A type of representation based on occupancy grids used to divide a space into discrete grid cells. For surface-type map in 3D this becomes voxels.
Grit-blasting	The abrasive removal of surface rust and/or paint using a high pressure grit stream.
Surface-type map	Model of the geometry and surface-type of surfaces in the environment.
RGB-D	The combination of a colour image represented in the RGB colour-space (red, green, blue) with the addition of depth data that corresponds with each colour image pixel.

Robot manipulator	In this thesis, this is a six-degree of freedom Denso industrial robotic manipulator, with a RGB-D sensor tool mounted on the end-effector.
Sensor package	Generally refers to an IR-based depth sensing camera, a colour camera and a light source.
Surface	The face of an object/structure in the environment.
Surface normal	A 3D vector perpendicular to a surface.
Surface-type	The appearance of a surface described by the colour and texture.
Textural appearance	The visual appearance of a surface that can be changed by the image capture conditions.
Viewpoint	A position in space and an orientation of a sensor that results from a manipulator pose $\vec{Q}$. This can also be expressed in terms of the homogeneous transformation matrix, ${}^{0}T_{s}(\vec{Q})$
Voxel	Volumetric Pixel which represents a 3D cube-like volume in Euclidean space.