

Robust Number Plate Detection Algorithms

A Thesis Submitted for the Degree of
Doctor of Philosophy

By
Wenjing Jia

in

Faculty of Information Technology
UNIVERSITY OF TECHNOLOGY, SYDNEY
AUSTRALIA
SEPTEMBER 2006

© Copyright by Wenjing Jia, 2006

UNIVERSITY OF TECHNOLOGY, SYDNEY
FACULTY OF INFORMATION TECHNOLOGY

The undersigned hereby certify that he has read this thesis entitled
“**Robust Number Plate Detection Algorithms**” by **Wenjing Jia**
and that in his opinion it is fully adequate, in scope and in quality, as a
thesis for the degree of **Doctor of Philosophy**.

Dated: September 2006

Research Supervisor: _____
Xiangjian He

CERTIFICATE

Date: **September 2006**

Author: **Wenjing Jia**

Title: **Robust Number Plate Detection Algorithms**

Degree: **Ph.D.**

I certify that this thesis has not already been submitted for any degree and is not being submitted as part of candidature for any other degree.

I also certify that the thesis has been written by me and that any help that I have received in preparing this thesis, and all sources used, have been acknowledged in this thesis.

Production Note:

Signature removed prior to publication.

Signature of Author

Acknowledgements

I would like to take this opportunity to express my sincere gratitude to my supervisor, Xiangjian He, for his unswerving encouragement, advice, help and invaluable suggestions, and for giving me the opportunity to do a PhD. He is such a kind, generous, helpful and kindhearted supervisor and friend to me, that I feel really happy, comfortable and unconstrained with him during my PhD study. I owe my research achievements to his excellent supervision.

Great thanks are also due to my co-supervisor, Massimo Piccardi, for his valued suggestions and constant support. I am very grateful to my colleagues, Qiang Wu and Huaifeng Zhang, for numerous discussions which have given me tremendous confidence on generating new ideas, and for their contributions to this thesis. Sincere thanks are also due to Tom Hintz, whose generosity and kindness will influence me my whole life. My sincere thanks are also due to my fellow research students and the staff of the faculty, especially those people listed below for providing various assistance for the completion of this research work: Chengqi Zhang, Maolin Huang, Tony Jan, Indrawati Nataatmadja, Fengjie Wu, Quang Vinh Nguyen, Hatice Gunes, Richard Xu, Lihong Zheng, Huaqing Wang, etc. My special thanks to my husband, Wenyong Wei, for his steadfast understanding, love and patience.

This work was edited in accordance with Parts D, Language and Illustrations, and E, Completeness and Consistency, of the Australian Standards for Editing Practice by Dr Bruce Howarth, who has taught in the IT field. Hereon, I would like to express my sincere thankfulness to Bruce for his professional work that has greatly improved the quality of this thesis.

Last but not least, this thesis could not have been completed without the financial support from the Faculty of Information Technology. The work in this thesis is partly supported by the Australian Research Council (ARC) under Large ARC-Discovery Grant No. DP0451666. I appreciate the travel support for attending international conferences which I received from the Faculty of Information Technology and the Vice-Chancellor's Conference Fund.

To my beloved Grandfather, Fulin.

Table of Contents

Table of Contents	vi
List of Tables	x
List of Figures	xi
List of Acronyms	xv
Abstract	xvi
1 Introduction	1
1.1 Automatic Number Plate Recognition	3
1.1.1 System Model	5
1.1.2 Image Acquisition	7
1.1.3 Number Plate Detection	8
1.1.4 Character Isolation	8
1.1.5 Character Recognition	9
1.1.6 Summary	10
1.2 Main Methods for Number Plate Detection	10
1.2.1 Texture-Based Methods	11
1.2.2 Vertical Lines-Based Methods	11
1.2.3 Colour-Based Methods	12
1.2.4 Shape-Based Methods	13
1.3 Main Problems with Number Plate Detection	14
1.3.1 Detecting Colour Number Plates	14
1.3.2 Robust Detection under Various Conditions	15
1.3.3 Real-Time Detection	15
1.4 The Author's Contributions to the Field of NPD by This Thesis . . .	16
1.4.1 A Region-Based NPD Algorithm	16
1.4.2 A Fast and Robust Edge-Based NPD Algorithm	18

1.4.3	Histogram-Based Methods for Classifying Number Plate Colours	19
1.4.4	Applications on Hexagonal Image Structure	20
1.5	Organisation of This Thesis	20
2	A Region-Based NPD Algorithm	21
2.1	Mean Shift for Image Segmentation	24
2.1.1	Mean Shift	24
2.1.2	Mean Shift for Image Segmentation	29
2.1.3	Segmentation Results	32
2.2	Feature Extraction	34
2.2.1	Feature 1: Rectangularity	35
2.2.2	Feature 2: Aspect Ratio	36
2.2.3	Feature 3: Edge Density	37
2.3	Classification	37
2.3.1	Statistical Data of the Three Features	38
2.3.2	Mahalanobis Distance-Based Classification	41
2.4	Experiments	42
2.4.1	Experimental Setup	42
2.4.2	Experimental Results and Discussion	42
2.5	Summary	49
3	A Fast and Robust Edge-Based NPD Algorithm	50
3.1	Overview of the Algorithm	54
3.1.1	Training	56
3.1.2	Testing	59
3.2	Vertical Edge Detection	61
3.2.1	Why Use Vertical Edges?	62
3.2.2	Conventional Method of Detecting Vertical Edges	66
3.2.3	Proposed Method of Detecting Vertical Edges	71
3.3	Global and Local Features	75
3.3.1	Local Haar-Like Features	76
3.3.2	Global Statistical Features	78
3.3.3	Computing Features Rapidly	82
3.4	Learning Classification Functions	86
3.4.1	AdaBoost: an Overview	88
3.4.2	Selecting Weak Classifiers	89
3.4.3	Cascade of Classifiers	91
3.5	Experiments	93
3.5.1	Selection of Data Set	93
3.5.2	Selection of Scaling Factors	96

3.5.3	Experimental Results and Discussion	97
3.6	Summary	103
4	Histogram-Based Classification of Number Plate Colours	105
4.1	Gaussian-Weighted Histogram Intersection	108
4.1.1	Conventional Histogram Intersection	109
4.1.2	Problems with Existing HI Algorithms	110
4.1.3	Gaussian-Weighted Histogram Intersection	112
4.1.4	Experiments	117
4.2	Colour Edge Cooccurrence Histogram	121
4.2.1	Colour Edge Cooccurrence Histogram	124
4.2.2	Reducing the Number of Colours	128
4.2.3	Comparing CECH Histograms Using Extended GWHI	130
4.2.4	Experiments	133
4.3	Comparison of Histogram-Based Image Matching Methods	139
4.4	Summary	142
5	Image Matching on Hexagonal Structure	143
5.1	Hexagonal Structure for Image Processing	144
5.1.1	Advantages of Hexagonal Structure in Image Processing	145
5.1.2	Hexagonal Image Representation	150
5.1.3	Hexagonal Structure Addressing	154
5.1.4	Hexagonal Image Processing	157
5.2	Symmetric Colour Ratio Gradients	159
5.2.1	Conventional Colour Ratio Gradients	160
5.2.2	Problem with the Conventional CRG	162
5.2.3	Symmetric Colour Ratio Gradients	165
5.2.4	Similarity Measurement	168
5.2.5	Experiments	170
5.3	Summary	176
6	Conclusions and Future Work	178
6.1	Conclusions	178
6.1.1	Number Plate Detection	178
6.1.2	Histogram-Based Methods for Classifying Number Plate Colours	179
6.1.3	Application of Image Matching on Hexagonal Structure	181
6.2	Future Work	181
6.2.1	Towards a Successful ANPR	181
6.2.2	Other Applications of the Proposed Image Matching Methods	182
6.2.3	Applications of Hexagonal Structure	183

Appendices	184
A A New Virtual Hexagonal Structure	185
A.1 The Virtual Hexagonal Structure	186
A.1.1 Square Grids and Hexagonal Grids	186
A.1.2 Construction of Hexagonal Pixels	187
A.1.3 Locating Hexagonal Pixels	190
A.1.4 Conversion Between Image Structures	191
A.2 Basic Image Transformations on the Virtual Hexagonal Structure . .	193
A.2.1 Translation	193
A.2.2 Rotation	194
A.3 Experimental Results	196
A.4 Conclusions	199
B Edge Detection on Virtual Hexagonal Architecture	201
B.1 Edge Detection	203
B.1.1 Image Smoothing	203
B.1.2 Gradient Detection	206
B.1.3 Edge Refining	208
B.2 Experimental Results	209
B.3 Conclusions	210
Author's Selected Publications for the PhD	214
Bibliography	218

List of Tables

1.1	The author's contributions to the state of the art of NPD by this thesis.	17
2.1	The statistical mean values of the three features for three classes of regions	39
4.1	Time required to compute image matching using the four methods. .	141

List of Figures

1.1	The general model of ANPR system.	5
1.2	An example of ANPR system.	6
2.1	The flow chart of the proposed region-based NPD algorithm.	23
2.2	Two examples of image segmentation using the mean shift algorithm.	33
2.3	Two typical examples of number plates that do not meet the assumption.	33
2.4	The estimated PDFs of the features of two kinds of number plates.	40
2.5	Interference characters (“ISUZU”) on a truck and the detection result.	44
2.6	Interference characters (two green letter P’s) that are next to the number plate and the detection result.	44
2.7	Comparison of the number plate detection results obtained using the texture-based method and our region-based method.	46
2.8	Comparison of the number plate detection results obtained using the texture-based method and our region-based method.	46
2.9	Estimated PDF of the error rate ϵ	47
3.1	The working flow of a cascade classifier.	55
3.2	The flow chart illustrating the principle of our new cascade classifier.	59
3.3	The smallest and biggest number plates that can be detected using our algorithm.	60
3.4	Original number plate images, their grey-level images and their gradient strength maps.	63
3.5	Examples of horizontally and vertically mounted number plates.	65

3.6	Various edge maps of a vehicle image.	67
3.7	A 3×3 region of an image (the z 's are grey-level values) and the Sobel masks used to compute the gradient at the centre point in x (horizontal) direction and y (vertical) direction respectively.	69
3.8	A non-vertical edge pixel in a 3×3 region of an image.	70
3.9	Comparison of the vertical edge maps of number plate images computed using the conventional method and our method.	72
3.10	Comparison of the vertical edge maps of a vehicle image obtained using the conventional method and our method.	74
3.11	Four types of Haar-like features shown relative to the scanning window.	76
3.12	Two regions that have typically different edge distributions.	79
3.13	The distributions of the vertical edges of the two regions shown in Figure 3.12.	80
3.14	The value of the integral image at pixel (x, y) is the sum of all the pixels which are located above and to the left of the current pixel.	83
3.15	Computing rectangular sum rapidly using the integral image.	84
3.16	Computing four types of Haar-like features using the integral image.	85
3.17	Examples of the number plates to be tested in our experiments.	94
3.18	The relationships between the scaling factor and the detection rate and average processing time respectively.	97
3.19	Examples of NPD results with normal conditions.	98
3.20	Examples of NPD results with strong illumination.	99
3.21	Examples of NPD results with interference characters.	100
3.22	Examples of NPD results with multiple number plates in one image.	101
3.23	Examples of NPD results with large view angle or blurry images.	102
3.24	Two unsuccessful detection cases.	103
4.1	Histogram intersection between a model image and a target image.	110
4.2	An example of the problem with the conventional HI algorithm.	111
4.3	The weight functions in the three HI algorithms.	115

4.4	The yellow model image used in our experiments.	119
4.5	Examples of yellow and white target images used in experiments. . .	119
4.6	Histogram matching results obtained using the three HI algorithms. .	120
4.7	Absolute positions in images where the object has been rotated. . . .	123
4.8	Computing the CECHE histogram at an edge point.	126
4.9	Number plate images and their quantised counterparts using the optimal colour quantisation method and the colour naming method. . . .	131
4.10	The histogram matching results using the proposed CECHE method. .	134
4.11	Image matching results using three different model images with 100 white target images.	136
4.12	Image matching results with tilted yellow images.	137
4.13	Image matching results with tilted white images.	137
4.14	Histogram matching results obtained using the four HI algorithms. . .	140
5.1	Connectivity on square grids and hexagonal grids.	147
5.2	The neighbourhood relationship on square grids and hexagonal grids.	148
5.3	The distance between pixels on square grids and hexagonal grids. . .	149
5.4	Reflection symmetries on square grids and hexagonal grids.	150
5.5	Using half-pixel shifted square pixels to represent hexagonal structure.	151
5.6	Rectangular pixels on a hexagonal sampling grid.	152
5.7	Simulated hyper pixel.	152
5.8	A cluster of seven mimic hexagons.	153
5.9	Image processing on the virtual Spiral Architecture.	154
5.10	The 2-axis oblique coordinate system for hexagonal structure.	155
5.11	Symmetrical hexagonal frame $\ast R3$	156
5.12	The one-dimensional spiral addressing scheme.	157
5.13	Locations of four neighbours that are involved in the computation of the colour ratio gradients at the central point.	161
5.14	Examples of the colour ratio gradients of a square image when the object is rotated in the image.	164

5.15	M_x and M_y in the conventional CRG.	165
5.16	Symmetric colour ratios M_1 , M_2 and M_3 at the central hexagonal pixel p and the locations of the six directly connected neighbours.	166
5.17	Example of the symmetric colour ratio gradients of a hexagonal image when the object is rotated in the image.	169
5.18	A vehicle image and its representation on virtual Spiral Architecture.	171
5.19	Histogram matching results obtained using the CRG and our SCRG algorithms.	173
5.20	An example of cutting a small part of a number plate image to form model images.	174
5.21	Similarity measurement with rotated model images.	176
A.1	The square lattice and the hexagonal lattice.	187
A.2	A single hexagonal pixel in the virtual hexagonal structure.	188
A.3	A cluster of seven constructed virtual hexagonal pixels.	189
A.4	An image of size 6×6 pixels tiled by the virtual hexagonal pixels. . .	192
A.5	Regions, layers and positions of hexagonal pixels	195
A.6	An example of image translation on the virtual hexagonal structure. .	197
A.7	An example of image rotation on the virtual hexagonal structure. . .	198
B.1	A 7-pixel neighbourhood in hexagonal structure.	206
B.2	The Sobel masks in three directions on hexagonal structure.	207
B.3	The three unit vectors indicated as the bold angled lines.	208
B.4	The original Lena image tested in our experiments.	209
B.5	Images processed by different filters on different image structures. . .	211
B.6	Edge maps of the filtered images shown in Figure B.5.	212

List of Acronyms

ANPR	Automatic Number Plate Recognition
CCH	Colour Cooccurrence Histogram
CCV	Colour Coherence Vector
CECH	Colour Edge Cooccurrence Histogram
CRG	Colour Ratio Gradient
EDV	Edge Density Variance
FFT	Fast Fourier Transform
GWHI	Gaussian-Weighted Histogram Intersection
HI	Histogram Intersection
ITS	Intelligent Transportation System
MER	Minimum Enclosing Rectangle
MPHM	Merged-Palette Histogram Matching
NPD	Number Plate Detection
OCR	Optical Character Recognition
PDF	Probability Density Function
ROI	Region of Interest
SA	Spiral Architecture
SCRG	Symmetric Colour Ratio Gradient

Abstract

Automatic Number Plate Recognition (ANPR) is an important Intelligent Transportation System (ITS) technology, which distinguishes each vehicle as unique by recognising the characters in their number plates via image analysis and pattern recognition techniques. In an ANPR system, the most crucial part is number plate detection. The research presented in this thesis focuses on the detection mechanism and will rely on a third-party Optical Character Recognition (OCR) software for character recognition.

Number Plate Detection (NPD) is a well-explored problem with many successful solutions. Although most of these solutions are reasonably fast and robust, they can be further improved to make them even faster and more robust to deal with various complex conditions in real-time.

This thesis first presents a region-based NPD algorithm, which provides much more accurate detection results than previous NPD algorithms and is robust against interference characters in images. Then, a fast and robust edge-based NPD algorithm is developed. This algorithm can detect various number plates under various conditions in real-time with a high detection rate and a very low false positive rate. Similar work has not been reported elsewhere.

Besides character information, the colour information of number plates also plays an important role in identifying each number plate as unique. Hence, this thesis also develops algorithms for classifying number plate colours. Histogram-based image matching methods are investigated, and a Gaussian Weighted Histogram Intersection (GWHI) algorithm is presented. This algorithm is shown to be much more robust

against various colour variations than previous methods. Furthermore, a novel Colour Edge Co-occurrence Histogram (CECH) method is presented. It is shown to be particularly applicable for rapidly matching compound objects, such as number plates.

Finally, histogram-based image matching technique on a hexagonal image structure is investigated. Gevers' idea of using Colour Ratio Gradient (CRG) for robust object matching is redefined on hexagonal structure, and a novel Symmetric Colour Ratio Gradient (SCRG) method is developed. Experimental results demonstrate that the proposed SCRG method outperforms the Gevers' CRG method.

More contributions can be found in the appendices. A new virtual hexagonal structure is proposed, on which the time used for mapping a square-based image to hexagon-based image is dramatically reduced. Two basic image transformation operations and a novel edge detection algorithm are performed on the new structure.

The results obtained in this thesis can also be applied to many other areas such as Character Detection, Text Detection, and Image/Video Retrieval.