

2009

## Illumination invariant face detection

Alister Cordiner  
*University of Wollongong*

Follow this and additional works at: <https://ro.uow.edu.au/theses>

**University of Wollongong**

**Copyright Warning**

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following: This work is copyright. Apart from any use permitted under the Copyright Act 1968, no part of this work may be reproduced by any process, nor may any other exclusive right be exercised, without the permission of the author. Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material.

Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

Unless otherwise indicated, the views expressed in this thesis are those of the author and do not necessarily represent the views of the University of Wollongong.

---

### Recommended Citation

Cordiner, Alister, Illumination invariant face detection, MComSc thesis, School of Computer Science and Software Engineering, University of Wollongong, 2009. <http://ro.uow.edu.au/theses/864>

## **NOTE**

This online version of the thesis may have different page formatting and pagination from the paper copy held in the University of Wollongong Library.

## **UNIVERSITY OF WOLLONGONG**

### **COPYRIGHT WARNING**

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site. You are reminded of the following:

Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

# Illumination Invariant Face Detection

A thesis submitted in fulfillment of the  
requirements for the award of the degree

**Master of Computer Science**

from

UNIVERSITY OF WOLLONGONG

by

**Alistair Cordiner**

School of Computer Science and Software Engineering

July 2009

© Copyright 2009

by

Alister Cordiner

All Rights Reserved

*Dedicated to*  
*Leonard and Sylvia*

# Declaration

This is to certify that the work reported in this thesis was done by the author, unless specified otherwise, and that no part of it has been submitted in a thesis to any other university or similar institution.

---

Alistair Cordiner  
6th July 2009

# Abstract

---

The purpose of face detection is to process input images in order to determine the locations of any faces in the image. Faces are complex objects and detecting them remains a challenging task for computer vision systems, despite the relative ease with which humans are able to do so. One of the major difficulties faced by face detection systems is challenging illumination conditions, such as low level lighting and cast shadows. This thesis reviews the state of the art face detection methods (with particular emphasis on the method of Viola and Jones) and explores methods of overcoming adverse illumination conditions. These methods can be broadly classified as invariant features, normalisation and variation modelling. Four novel approaches to overcoming illumination that fall into these 3 categories are proposed in this thesis, namely: (i) log-ratio Haar-like features; (ii) DC Haar-like features; (iii) local variance normalisation; and (iv) classifier fusion. Furthermore, a new type of feature called the *generalised integral image feature* (GIIF) is proposed as an alternative to Haar-like features. The GIIF method is not specifically related to illumination invariant face detection, but instead applies to the more general task of face detection and is therefore presented as a separate chapter. Experimental results on standard face databases are provided for all of the proposed methods to verify that they achieve improved accuracy.

# Acknowledgements

---

First and foremost, I would like to thank my supervisors, Professor Philip Ogunbona and Dr Wanqing Li. Without their wisdom, guidance, constructive feedback and encouragement this work would not have been possible. I wish to thank all of my student colleagues, particularly the members of the Advanced Multimedia Research Lab at the University of Wollongong, for the opportunities to exchange ideas and share diverse points of view at the group meetings. I also wish to thank my employers, particularly Dr Tarik Hammadou, for being flexible and allowing me to complete my thesis while being employed. And finally, I would like to thank my family, especially my parents, Leonard and Sylvia, for being supportive and encouraging over these past two years. To them I dedicate this thesis.



# Contents

---

<b>Abstract</b>	<b>v</b>
<b>Acknowledgements</b>	<b>vi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Overview . . . . .	2
1.2 Contributions of the thesis . . . . .	2
1.3 Publication list . . . . .	3
1.4 Organisation of the thesis . . . . .	3
<b>2 Literature review</b>	<b>5</b>
2.1 Overview . . . . .	6
2.2 Face detection methods . . . . .	7
2.2.1 Knowledge-based methods . . . . .	11
2.2.2 Feature-based methods . . . . .	12
2.2.3 Template matching methods . . . . .	13
2.2.4 Appearance-based methods . . . . .	14
2.2.5 Seminal papers . . . . .	17
2.2.6 Face databases . . . . .	18
2.2.7 Evaluating face detection systems . . . . .	20
2.3 Challenges in face detection . . . . .	22
2.3.1 Low resolution or poor quality images . . . . .	22
2.3.2 Occlusion by foreign objects . . . . .	23
2.3.3 Illumination variations . . . . .	25
2.3.4 Facial expressions and structural obstructions . . . . .	26
2.3.5 Pose (rotation) . . . . .	27
2.3.6 Artificial faces . . . . .	29
2.4 Evaluation of Viola and Jones face detection . . . . .	30

2.5	Discussion . . . . .	35
<b>3</b>	<b>Illumination invariant face detection</b>	<b>36</b>
3.1	Theory of light . . . . .	37
3.1.1	Symbols and notation . . . . .	37
3.1.2	Overview . . . . .	37
3.1.3	Lighting environments . . . . .	45
3.1.4	Effects of illumination . . . . .	46
3.1.5	Human perception . . . . .	47
3.1.6	Approaches to illumination invariance . . . . .	47
3.2	Invariant features . . . . .	48
3.2.1	Intensity derivative (gradient) filters . . . . .	50
3.2.2	DCT-based features . . . . .	51
3.2.3	Gabor wavelet filters . . . . .	51
3.2.4	Haar-like wavelet filters . . . . .	52
3.2.5	Experiments with invariant features . . . . .	52
3.2.6	Discussion . . . . .	54
3.3	Normalisation and relighting . . . . .	54
3.3.1	Monotonic transfer functions . . . . .	56
3.3.2	Homomorphic filtering . . . . .	58
3.3.3	Basis sets . . . . .	59
3.3.4	Ratio images . . . . .	63
3.3.5	Experiments with normalisation . . . . .	66
3.3.6	Discussion . . . . .	68
3.4	Variation modelling . . . . .	68
3.4.1	Illumination cones . . . . .	71
3.4.2	3-D linear subspace . . . . .	72
3.4.3	Principal Component Analysis (PCA) . . . . .	73
3.4.4	Independent Component Analysis (ICA) . . . . .	73
3.4.5	Fisher's Linear Discriminant (FLD) . . . . .	74
3.4.6	Non-negative Matrix Factorisation (NMF) . . . . .	74
3.4.7	Discussion . . . . .	76
<b>4</b>	<b>Proposed methods</b>	<b>78</b>
4.1	Invariant features . . . . .	79
4.1.1	Ratio features . . . . .	79

4.1.2	DC features . . . . .	80
4.1.3	Experimental results . . . . .	82
4.1.4	Discussion . . . . .	84
4.2	Normalisation/relighting approaches . . . . .	85
4.2.1	Local variance normalisation . . . . .	85
4.2.2	Experimental results . . . . .	87
4.2.3	Discussion . . . . .	88
4.3	Variation modelling . . . . .	88
4.3.1	Clustering and classifier fusion . . . . .	89
4.3.2	Experimental results . . . . .	95
4.3.3	Discussion . . . . .	98
4.4	Comparison of methods . . . . .	99
4.4.1	Experimental results . . . . .	99
4.4.2	Example face detection results . . . . .	100
4.4.3	Discussion . . . . .	102
<b>5</b>	<b>Generalised integral image features</b>	<b>105</b>
5.1	Background . . . . .	106
5.1.1	Integral images . . . . .	106
5.1.2	Feature selection . . . . .	107
5.1.3	Genetic algorithms . . . . .	108
5.2	Related work . . . . .	110
5.3	Method . . . . .	112
5.3.1	Features . . . . .	112
5.3.2	Training . . . . .	115
5.3.3	Translation invariant features . . . . .	117
5.3.4	Computational complexity . . . . .	121
5.3.5	Practical considerations . . . . .	121
5.4	Experiments and results . . . . .	121
5.4.1	Varying the mutation and crossover rates . . . . .	121
5.4.2	Varying the value of $\alpha$ . . . . .	122
5.4.3	Varying the range of $\mathbf{g}$ . . . . .	123
5.4.4	Comparison to Viola and Jones . . . . .	125
5.5	Discussion . . . . .	125
<b>6</b>	<b>Conclusion</b>	<b>127</b>

<b>A</b>	<b>Viola and Jones face detection</b>	<b>131</b>
A.1	Haar-like features . . . . .	132
A.2	AdaBoost . . . . .	134
A.3	Cascade generation . . . . .	138
A.4	Merging algorithm . . . . .	139
A.5	Extensions to Viola and Jones . . . . .	139
A.5.1	AdaBoost algorithm . . . . .	139
A.5.2	Weak learners . . . . .	140
A.5.3	Haar-like features . . . . .	141
	<b>Bibliography</b>	<b>142</b>

# List of Tables

---

2.1	Commonly used face databases . . . . .	19
2.2	Average detection rates at original scale . . . . .	31
2.3	Average number of false positives at original scale . . . . .	32
4.1	Comparison of face detection results . . . . .	101
4.2	Examples of false positives and false negatives for all illumination invariant approaches	103
5.1	Comparison of AUC and error rate . . . . .	125

# List of Figures

---

2.1	Contoured sketch of a human face [45]	7
2.2	Face detection system topology	7
2.3	Combining results from multi-resolution, multi-face detection	8
2.4	Pattern classification approach to detection	9
2.5	Knowledge-based architecture	11
2.6	Feature-based architecture	12
2.7	Template matching architecture	13
2.8	Example of AAM iterative deforming [37]	13
2.9	Appearance-based architecture	14
2.10	Example images from the AR face database [113]	20
2.11	The importance of context	23
2.12	Examples of occlusion by sunglasses and a scarf [113]	24
2.13	Examples of different facial expressions [108]	26
2.14	Examples of a frontal face and out-of-plane rotated faces [9]	27
2.15	The four main approaches to rotation-invariant face detection.	28
2.16	Examples of artificial faces. [8]	30
2.17	Face size vs. detection rate	33
2.18	Average intensity vs. detection rate	33
2.19	Examples of images with low and high average face intensity which were not detected	34
2.20	Examples of images where the face detector performed poorly	34
3.1	Specular vs. diffuse reflection	38
3.2	Different types of reflection from the skin [44]	39
3.3	Scattering effect of diffuse surfaces “blurs” the incident light as it is reflected	40
3.4	Histogram skewness indicates the type of reflection	41
3.5	Example global and individual geometric models generated from 3-D face data [27]	41

3.6	An image can be represented as the pixel-wise product of a reflectance image and an illuminance image . . . . .	43
3.7	Attached and cast shadows from an incident light ray [23] . . . . .	44
3.8	Cube and spherical environment maps showing some possible incident light directions	45
3.9	Examples of varying illumination [59] . . . . .	46
3.10	Edge maps are unstable under strong illumination variations and the computed edges do not necessarily correspond to changes in reflectance . . . . .	49
3.11	Gabor wavelet filters of varying frequencies and orientations . . . . .	51
3.12	Frequency response of the Haar-like features proposed by Viola and Jones . . . . .	52
3.13	Examples of invariant feature values, where the top left image is the original input image from the Bio ID database . . . . .	53
3.14	Performance of a Viola and Jones classifier using different invariant features . . . . .	55
3.15	Homomorphic filtering using the discrete Fourier transform (DFT) . . . . .	58
3.16	Example of homomorphic filtering applied to a face . . . . .	59
3.17	Decomposition of an input image sequence into reflectance and illumination images using the method of [174] . . . . .	60
3.18	Normalisation methods . . . . .	67
3.19	Mean and variance normalisation . . . . .	69
3.20	Visual representation of the illumination cone of a face [2] . . . . .	72
3.21	(a) NMF illumination basis images. (b) Relighting results (top row are the input images, bottom row are the relit images). . . . .	75
3.22	DCT, PCA (“eigenfaces”) and NMF basis images generated for the experiments . . . .	77
4.1	Haar-like features $H_1$ to $H_4$ were used by Viola and Jones. Feature $H_5$ is the proposed DC feature. . . . .	81
4.2	Example images where the classifier fails ranked in order of confidence weighting from top left to bottom right . . . . .	81
4.3	Comparison of the error rate of face detectors trained with a single Haar-like feature. Feature $H_5$ is the proposed DC feature. . . . .	82
4.4	Comparison of the classification margin $g_i$ versus illumination angle on the Yale B database for (a) the DC features and (b) the ratio features. . . . .	83
4.5	Comparison of error rate on the Yale B database. . . . .	84
4.6	False positive (a) before and (b) after variance and mean normalisation is applied. . .	85
4.7	Comparison of different types of variance normalisation . . . . .	86

4.8	Comparison of the classification margin $g_i$ versus illumination angle on the Yale B database. . . . .	87
4.9	Comparison of error rate on the Yale B database. . . . .	88
4.10	Examples of 3 possible illumination classes . . . . .	89
4.11	High-level block diagram of the proposed face detection system. . . . .	89
4.12	Examples of faces clustered into face illumination classes using $k$ -means clustering . .	92
4.13	Example membership matrices for crisp and fuzzy clustering . . . . .	92
4.14	Calculating the output of monolithic and multiple classifiers . . . . .	93
4.15	The first 10 Haar-like features selected for the (a) monolithic and (b) clustered face detectors. . . . .	96
4.16	The ROC curve shows that all of the multiple classifier variants outperform the monolithic classifier method. . . . .	97
4.17	Confusion matrices showing the detection rates across the different face illumination classes. . . . .	98
4.18	Comparison of ROC curves . . . . .	101
5.1	Input image and its corresponding integral image . . . . .	107
5.2	Region A can be computed as $y_1 + y_4 - (y_2 + y_3)$ [48] . . . . .	108
5.3	Example genetic algorithm [51] . . . . .	109
5.4	Some examples of Haar-like features used to represent $\mathbf{h}$ . . . . .	110
5.5	Example integral image look-ups used to calculate a single Haar-like feature value. . .	113
5.6	Example integral image look-ups used to calculate a single generalised integral image feature value. . . . .	113
5.7	Approximation of a Viola and Jones Haar-like feature in the GIIF feature space constrained by $\mathbf{g} \in \{-1, 0, +1\}$ . . . . .	115
5.8	Calculation of an “offset” integral image . . . . .	118
5.9	Example conversions from a GIIF $g$ to a translation invariant GIIF $\tilde{g}$ . . . . .	120
5.10	Effect of varying the mutation and crossover rates . . . . .	122
5.11	Effect of varying the $\alpha$ value on the error rate and sparsity . . . . .	122
5.12	“Equivalent” Haar-like features $\mathbf{h}$ generated by varying the value of $\alpha$ from a high value (left) to a low value (right) . . . . .	124
5.13	Effect of varying the range of $\mathbf{g}$ . . . . .	124
5.14	Comparison of the ROC curves . . . . .	126
A.1	Haar-like features originally proposed by [170], where the darker rectangles are the positive regions and lighter rectangles are the negative regions . . . . .	132



A.2	Example of an integral image . . . . .	133
A.3	AdaBoost viewed as a neural network . . . . .	135
A.4	Weak classifiers with lower training errors are given higher weightings . . . . .	137
A.5	Structure of the face detector cascade . . . . .	138
A.6	Examples of multiple detections before merging is applied [142] . . . . .	139