



Kent Academic Repository

**Alsufyani, Hamed (2018) *Skin Texture as a Source of Biometric Information*.
Doctor of Philosophy (PhD) thesis, University of Kent,.**

Downloaded from

<https://kar.kent.ac.uk/72905/> The University of Kent's Academic Repository KAR

The version of record is available from

This document version

UNSPECIFIED

DOI for this version

Licence for this version

CC BY (Attribution)

Additional information

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in *Title of Journal*, Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).

Skin Texture as a Source of Biometric Information

A Thesis Submitted to the University of Kent
For the Degree of Doctor of Philosophy
In Electronic Engineering

By
Hamed Alsufyani
October 2018

© Hamed Alsufyani, Canterbury, UK, 2018

Abstract

Traditional face recognition systems have achieved remarkable performances when the whole face image is available. However, recognising people from partial view of their facial image is a challenging task. Face recognition systems' performances may also be degraded due to low resolution image quality. These limitations can restrict the practicality of such systems in real-world scenarios such as surveillance, and forensic applications. Therefore, there is a need to identify people from whatever information is available and one of the possible approaches would be to use the texture information from available facial skin regions for the biometric identification of individuals.

This thesis presents the design, implementation and experimental evaluation of an automated skin-based biometric framework. The proposed system exploits the skin information from facial regions for person recognition. Such a system is applicable where only a partial view of a face is captured by imaging devices. The system automatically detects the regions of interest by using a set of facial landmarks. Four regions were investigated in this study: forehead, right cheek, left cheek, and chin. A skin purity assessment scheme determines whether the region of interest contains enough skin pixels for biometric analysis. Texture features were extracted from non-overlapping sub-regions and categorised using a number of classification schemes. To further improve the reliability of the system, the study also investigated various techniques to deal with the challenge where the face images may be acquired at different resolutions to that available at the time of enrolment or sub-regions themselves be partially occluded. The study also presented an adaptive scheme for exploiting the available information from the corrupt regions of interest.

Extensive experiments were conducted using publicly available databases to evaluate both the performance of the prototype system and the adaptive framework for different operational conditions, such as level of occlusion and mixture of different resolution skin images. Results suggest that skin information can provide useful discriminative characteristics for individual identification. The comparison analyses with state-of-the-art methods show that the proposed system achieved a promising performance.

Dedication

- To my beloved mother Magbola.. Without her sacrifice, supplication, support, encouragement during the period of my PhD study in UK, there would never have been any chance for this work to happen.
- To my wife Dr Nawal.. The light of my life.. Without her support, care, sincere prayers and encouragement.. I would never have done this work.
- To my sons Hatim and Malik.. The stars of my future.. They will be proud of their parents.
- To my brothers and sisters.. Many thanks to them for their support and encouragement.
- To my brother-in-law Dr Abdulmajeed.. I deeply appreciate his unlimited and incredible support.
- To my brothers-in-law.. I am really grateful to them for their help.

Thanks to all of you

Hamed..

Canterbury

30 September 2018

Acknowledgements

"رَبِّ اشْرَحْ لِي صَدْرِي ، وَبَسِّرْ لِي أَمْرِي ، وَاحْلُلْ عُقْدَةً مِّن لِّسَانِي ، يَفْقَهُوا قَوْلِي"

"Oh my Lord; lift up my heart, ease my duty, untie my tongue, so that my words are understood"

The Holy Quran (Surat Ta ha, verses 25-28)

All praise and thanks are due to the Almighty Allah who always guides me to the right path and has helped me to complete this thesis.

I wish to express my deepest gratitude to my supervisors, Prof Farzin Deravi and Dr Sanaul Hoque. They gave me this precious opportunity to pursue PhD under their supervision. They have profound expertise and broad research vision which have guided me to several fruitful directions for my PhD research. I am grateful for the advices, comments, and encouragements received from them during this study.

I am grateful to the University of Kent and its School of Engineering and Digital Arts for the effort they made. The University of Kent and its School of Engineering and Digital Arts do not only have an open and inclusive multicultural environment but also create an excellent study and research atmosphere for international students. The technical support is solid and the devices are significantly contributing to the completion of my PhD research study. I would like to especially thank the administration team and the reception team of the school led by Prof Farzin Deravi for their professional support during my PhD study.

Special thanks to my friend.. Dr Md Moinul Hossain who has kindly given me tremendous help and guidance for my research.

I would like to express appreciation to my fellow lab-mates in the Intelligent Interactions Group for the inspiring discussions and for all the fun we have had in the last four years.

Content

Abstract i

Acknowledgements.....iii

Content iv

List of Tables xi

List of Figures.....xiv

List of Abbreviations.....xvii

Publications and Dissemination.....xix

Chapter 1 Introduction..... 1

1.1 An Introduction to Biometric Recognition Systems 1

1.2 Skin Texture as a Biometrics 3

1.3 Motivation and Objectives of the Research 5

1.4 Contributions..... 6

1.5 Thesis Structure..... 8

Chapter 2 Literature Review 11

2.1	Introduction	11
2.2	Challenges for Face Recognition	11
2.3	Historical Developments in Face Recognition Systems	12
2.4	Human Skin.....	13
2.4.1	Skin Texture and Disease Detection	14
2.4.2	Skin Texture and Cosmetology	14
2.4.3	Skin Information for Personal Recognition	15
2.4.4	Skin texture biometrics in Low-Resolution Images.....	20
2.5	Generic Face Skin Biometric System	21
2.5.1	Face Detection.....	22
2.5.2	Face Normalisation	22
2.5.2.1	Geometric Normalisation	23
2.5.2.2	Photometric Normalisation	23
2.5.3	Feature Extraction and Classification	24
2.6	Databases.....	24
2.7	Summary	24
Chapter 3 Previous Work on Texture Analysis		27
3.1	Introduction	27
3.2	History of Texture Analysis	27
3.3	Review of Texture Analysis Definitions and Challenges	28
3.3.1	Texture Analysis	28
3.3.2	Issues Related to Texture Analysis	29
3.4	Feature Extraction from Texture.....	30
3.4.1	Co-Occurrence Features.....	31

3.4.2	Local Binary Pattern (LBP) Operator	31
3.4.2.1	A Taxonomy of the LBP Operator.....	32
3.4.2.2	Extension of LBP	33
3.4.2.3	Rotation Invariant of LBP Operator.....	34
3.4.2.4	Mappings of the LBP Labels: Uniform Patterns.....	34
3.4.2.5	Multi-Scale LBP	35
3.4.3	Gabor Filter-based Texture Features.....	36
3.4.3.1	1-D Fourier Transform.....	37
3.4.3.2	2-D Gabor filter.....	38
3.5	Feature Classification.....	40
3.5.1	Distance Metrics.....	40
3.5.2	<i>K</i> -Nearest Neighbour (<i>k</i> -NN).....	42
3.5.3	Support Vector Machines (SVM)	42
3.5.4	Sparse Representation Classifier (SRC).....	46
3.6	Biometric Information Fusion Strategies	48
3.6.1	Sensor-level Fusion.....	49
3.6.2	Feature-level Fusion.....	49
3.6.3	Score-level Fusion.....	49
3.6.4	Decision-level Fusion.....	50
3.7	Measuring Biometric System Performances	50
3.8	Summary	51
Chapter 4 Detecting Facial Skin Regions for Biometric Use		52
4.1	Introduction	52
4.2	Challenges for Skin Detection	53

4.3	Previous Works on Skin Detection	53
4.4	A General Framework for Skin Purity Analysis	55
4.5	Skin Colour Schemes	56
4.6	Experimental Setup	58
4.6.1	Databases.....	58
4.6.1.1	The Skin Segmentation Dataset	58
4.6.1.2	XM2VTS Database	60
4.6.2	Generating the Ground Truth for Purity Analysis.....	60
4.6.3	Automated Skin Purity Analysis	63
4.7	Experimental Results and Discussion	63
4.7.1	Performance of the Proposed Skin Pixel Detection	63
4.7.2	Performance of Skin Detection Regional Purity Analysis	65
4.7.3	Analysis of Computational Cost	68
4.8	Summary	68

Chapter 5 Exploring the Potential of Facial Skin Regions for Identity

	Information.....	70
5.1	Introduction	70
5.2	General Framework of the Proposed System.....	72
5.2.1	Detecting Facial Landmarks for ROI Segmentation	74
5.2.2	Geometric Skin Localisation Measurements	75
5.2.3	Skin Pixel Detection for Skin Purity Analysis	77
5.2.4	Feature Extraction	77
5.2.4.1	Local Binary Pattern (LBP)	77
5.2.4.2	Gabor Filters	78

5.2.4.3	Feature Representation for LBP and Gabor Filters.....	79
5.2.5	Skin Features Classification.....	79
5.3	Experimental Setup.....	80
5.4	Experimental Analysis.....	80
5.4.1	Facial Landmark Detection.....	80
5.4.2	Skin Purity Analysis.....	81
5.4.3	Performance of Skin Biometrics.....	83
5.4.4	Comparison with Existing Systems.....	86
5.5	Further Improvement using Multi Scheme Purity Assessment for Skin-based Biometrics.....	86
5.5.1	Improving the Skin Purity Analysis Assessment.....	87
5.6	Experimental Results and Discussion.....	88
5.6.1	Skin Region Purity Analysis.....	88
5.6.2	Biometric Performance Analysis.....	91
5.6.3	Comparative Analysis.....	95
5.7	Summary.....	97

Chapter 6 Usability of Skin Texture Biometrics for Low-Resolution

	Images.....	99
6.1	Motivation.....	100
6.2	Experimental Setup.....	101
6.2.1	Generation of Multi Scale Dataste.....	101
6.3	Experimental Results and Discussion.....	103
6.3.1	Experiment 1: The Effect of Resolution.....	104
6.3.2	Experiment 2: The Effect of Mismatched Resolution between the Enrolment and Verification.....	106

6.3.3	Experiment 3: Incorporating Rejections in the Decision Process	113
6.3.4	Experiment 4: Effect of Multi-scale Enrolment using Multiple Parameters	115
6.3.5	Experiment 5: Analysis of Fusion-based Scheme.....	120
6.3.6	The Performance Comparison between Experiments	122
6.4	Summary	124

Chapter 7 Adaptive Skin Extraction Features Using Biometric

	Information.....	125
7.1	Introduction	125
7.2	The Concept and the Motivation.....	125
7.3	Experiments.....	126
7.3.1	Experiment 1: Skin-based Biometric using Information from each Sub- region.....	127
7.3.2	Experiment 2: Adaptation of Skin Features based on Extracting Available Information of ROI.....	129
7.3.3	Experiment 3: Adaptive Skin Feature Scheme at Arbitrary Scenarios for Skin-based Recognition System.....	133
7.4	Summary	136

Chapter 8 Conclusions and Future Work 137

8.1	Conclusions	137
8.2	Summary of Findings	137
8.3	Recommendations for Future Work.....	140
8.3.1	Skin Detection Improvement	140
8.3.2	Improvement of Localising Facial Skin Regions.....	141

8.3.3 Improvement of Feature Extraction Technique 141
8.3.4 Improvement of Skin Biometric in Less Constrained Conditions 142

Bibliography 143

List of Tables

Table 2.1 State of the art n facial skin texture biometric performance	19
Table 2.2 Widely used face image databases	25
Table 4.1 Ground truth of skin and non-skin images of the four facial regions from XM2VTS database	61
Table 4.2 Some samples of ROIs graded as accepted and rejected by the evaluators (ground truth)	61
Table 4.3 Performance of skin/non-skin detection technique compared with other published results.....	64
Table 4.4 Skin detection framework performance of the three schemes for the four independent facial skin regions (the threshold value is set to 98%)	67
Table 4.5 Computational time for the three skin detection schemes.....	68
Table 5.1 ROI dimensions based on different face landmarks, where $d1$ is the distance between the two eyes, $d2$ is the length of the dorsum nasi, and $d3$ is the distance between the outer mouth corners	76
Table 5.2 Recognition accuracy (rank-one) at a defined threshold for the four ROIs (%)	84
Table 5.3 Proposed system's recognition performance compared to earlier methods using the same database (XM2VTS).....	86
Table 5.4 Population coverage of the three skin purity assessment schemes	91
Table 5.5 Rank-one recognition rates for forehead region using different skin purity assessment schemes (purity threshold = 98%).....	92
Table 5.6 Rank-one recognition rates for right cheek region using different skin purity assessment schemes (purity threshold = 98%).....	92
Table 5.7 Rank-one recognition rates for left cheek region using different skin purity assessment schemes (purity threshold = 98%).....	93

Table 5.8 Rank-one recognition rates for chin region using different skin purity assessment schemes (purity threshold = 98%).....	93
Table 5.9 Comparison with other published schemes using the XM2VTS database.....	96
Table 6.1 ROI size ranges after bicubic interpolation using different scaling factors .	103
Table 6.2 Recognition rates for facial skin regions at different resolutions.....	105
Table 6.3 Best recognition rates with different R values for training and test sets when matching images at different scales	108
Table 6.4 Recognition rates with different training and test set R values for matching different right cheek resolutions.....	109
Table 6.5 Best recognition rates with different training and test set R values for matching right cheek skin images of different scales (summarised from Table 6.4)	110
Table 6.6 Recognition rates with different training and test set R values for matching left cheek skin images of different resolutions	111
Table 6.7 Best recognition rates with different training and test set R values for matching left cheek images of different scales (summarised from Table 6.6)	111
Table 6.8 Recognition rates with different training and test set R values for matching chin skin images of different resolutions	112
Table 6.9 Best recognition rates with different training and test set R values for matching chin images at different scales (summarised from Table 6.8).....	112
Table 6.10 Forehead region recognition rates (rank-one) for test set images of unknown size based on different combination techniques for training and test sets ...	117
Table 6.11 Right cheek recognition rates (rank-one) for test set images of unknown size based on different combination techniques for training and test set.....	118
Table 6.12 Left cheek recognition rates (rank-one) for test set images of unknown size based on different combination techniques for training and test set.....	119

Table 6.13 Chin recognition rates (rank-one) for test set images of unknown size based on different combination techniques for training and test set.....	119
Table 6.14 Recognition rates (rank-one) for test set images of unknown size based on score fusion techniques using k -NN classifier.....	121
Table 6.15 Results of comparison between Experiment 1 and Experiment 2 for recognition rates	123
Table 7.1 Recognition rates of single patch obtained from XM2VTS database	128
Table 7.2 Recognition rates (rank-one) for forehead region based on LBP descriptor using 1NN and SVM classifiers.....	131
Table 7.3 Recognition rates using part of ROI available for feature extraction.....	135

List of Figures

Figure 1.1 Schematic diagram of a skin biometric system	4
Figure 1.2 The outline of the thesis	10
Figure 2.1 Some significant milestones in the history of face recognition systems [2].	13
Figure 2.2 Fields of science and technology utilising skin appearance [24]	14
Figure 3.1 Timeline of work on texture representation [89]	28
Figure 3.2 Example of the original LBP operator [84]	33
Figure 3.3 Example of the basic LBP operator using different P and R : circular, (a) with (8,1), (b) with (16,2), and (c) with (24,3) neighbourhoods [114]	33
Figure 3.4 Example of MB-LBP showing (a) the original LBP, and (b) the 9×9 MB-LBP descriptor. For each sub-region, the average sum of image intensity is computed and thresholded by the centre block to attain MB-LBP [115]	36
Figure 3.5 Evolution of the LBP descriptor [89]	36
Figure 3.6 Separation of the data into two classes	44
Figure 4.1 General structure of the proposed skin detection framework	55
Figure 4.2 Proposal for a general skin detection system. The RGB skin image input is transformed to HSV and YCbCr. While skin modelling Scheme1 uses only one colour space (RGB), skin modelling Scheme2 combines two further colour spaces (HSV and YCbCr). Skin-modelling Scheme2 is based on H, Cb and Cr values.....	59
Figure 4.3 Manual labelling of the four facial skin regions by 15 observers to create a ground truth from the XM2VTS database	62
Figure 4.4 ROC curves of thresholding optimisation for (a) forehead, (b) left cheek (c) right cheek, (d) chin skin region using the three different schemes.....	66
Figure 5.1 Framework of the proposed skin-based biometric system.....	73
Figure 5.2 Facial landmarks detected by Chehra Face Tracker	74

Figure 5.3 Geometric facial landmark measurements for ROI localisation. These measurements were empirically determined using Chehra software. All ROIs fall within the face boundary	76
Figure 5.4 Some examples of users (a) whose facial landmarks were detected and (b) were not detected by Chehra Face Tracker software	81
Figure 5.5 Flowchart of the proposed purity assessment technique using skin pixel detection Scheme 2	81
Figure 5.6 Examples of accepted and rejected ROIs based on the skin purity assessment technique	82
Figure 5.7 Subjects with useable regions by set of thresholds for forehead and chin region.....	82
Figure 5.8 CMC curves of recognition rates for the two ROIs (forehead and chin) based on LBP operator using k -NN and SVM classifiers	85
Figure 5.9 Recognition rates for (a) forehead and (b) chin region based on LBP operator using two different classifiers (k -NN and SVM) with different thresholds	85
Figure 5.10 (a) TPR and (b) TNR of purity assessment schemes for various threshold values of forehead region	89
Figure 5.11 (a) TPR and (b) TNR of purity assessment schemes for various threshold values of chin region	89
Figure 5.12 HTER of purity assessment scheme 3 for various threshold values of facial skin regions	90
Figure 5.13 CMC curves when using (a) the LBP descriptor and (b) Gabor wavelets with the SVM classifier.....	95
Figure 6.1 Examples of facial skin images and skin regions of interest at different scales	102
Figure 6.2 Schematic of image scaling framework.....	105

Figure 6.3 Recognition rates with different training and test set R values for matching different resolutions of forehead skin images. The protocol was to train Dataset A and test (a) Dataset B and (b) Dataset C	107
Figure 6.4 Recognition rates with different training and test set R values for matching different resolutions of forehead skin images. The protocol was to train Dataset B and test (a) Dataset A and (b) Dataset C	107
Figure 6.5 Recognition rates with different training and test set R values for matching different resolutions of forehead skin images. The protocol was to train Dataset C and test (a) Dataset A and (b) Dataset B	107
Figure 6.6 Flowchart of error-rejection technique	113
Figure 6.7 Rejection rates for (a) forehead, (b) right cheek, (c) left cheek, and (d) chin region at different resolutions	114
Figure 6.8 Flowchart of the proposed technique for matching different image scales	116
Figure 6.9 CMC curves for the three training set combinations (A, B, and C) with all facial skin regions with $R=4$	120
Figure 6.10 Schematic of system framework for score fusion technique	121
Figure 6.11 CMC curve for the proposed system's recognition rate using max rule ..	122
Figure 7.1 Recognition rates for a single sub-region (patch) obtained from forehead skin	129
Figure 7.2 Schematic diagram for adaptive skin biometric system	130
Figure 7.3 CMC curves of rank-n recognition rates of different sub-regions available for feature extraction	132

List of Abbreviations

Abbreviation	Meaning
CMC	cumulative match characteristic
ROC	receiver operating characteristic
FA	False Acceptance
FR	False Rejection
FAR	False Acceptance Rate
FRR	False Rejection Rate
EER	equal error rate
HTER	Half total error rates
AUC	area under the curve
GAR	genuine acceptance rate
TER	total error rate
ROI	region of interest
ASM _s	Active Shape Models
AAM _s	Active Appearance Models
LBP	Local Binary Pattern
SVM _s	Support Vector Machines
Dpi	dots per inch
PLS	Partial least squares regression
SIFT	Scale Invariant Feature Transform
MLBP	Multiscale Local Binary Patterns
LDA	linear discriminant analysis
HEC	hierarchical ensemble classifier
LGFV	local Gabor feature vector
PCA	Principal Component Analysis
NPD	Normalized Pixel Difference
GLCM	Grey Level Co-occurrence Matrix
MB-LBP	Multi-scale Block Local Binary Pattern
<i>k</i> -NN	<i>k</i> -Nearest Neighbour

RBF	Radial basis function
SRC	sparse representation classifier
FR	Facial recognition
LoG	Laplacian of Gaussian

Publications and Dissemination

The following publications were produced during the course of the work leading to the preparation of this thesis.

1. H. Alsufyani, S. Hoque, and F. Deravi, "Automated skin region quality assessment for texture-based biometrics," in *Seventh International Conference on, Emerging Security Technologies (EST)*, 2017, pp. 169-174.
2. H. Alsufyani, S. Hoque, and F. Deravi, "Exploring the Potential of Facial Skin Regions for the Provision of Identity Information," in *Proceeding of the 7th IET International Conference on Imaging for Crime Detection and Prevention (ICDP-16)*, Madrid, Spain, 2016, pp. 1-6.

Chapter 1

Introduction

1.1 An Introduction to Biometric Recognition Systems

Identity management is crucial in ensuring adequate security in both the public and private sectors. In the early years of the digital age, a person's identity was confirmed using traditional methods such as a secret text string (a password) or a physical object such as a token or card, or by some combination of both. Such security measures may fail to verify the user when a password is forgotten or a card is lost. It is clear that while complex passwords are hard to be memorised, simple or short passwords can be easily guessed/stolen or even shared. Additionally, traditional mechanisms do not provide strong evidence for post-event person recognition such as suspect identification at crime scene.

Technological advances have created a need for more secure methods that overcome the limitations of traditional approaches. Biometric technologies have great potential in this context. Biometric systems offer a neutral and more reliable solution to the problem of person recognition. Biometric recognition originally refers to the automated verification/identification of an individual's identity based on measurements of their physical or behavioural traits. Such systems have become commonplace for secure access control. Since biometric characteristics are inherent to individuals, it is difficult

to be manipulated, shared, or forgotten. However, a key concern is which biological measurements are suitable as biometric characteristics. To answer that question, biometric characteristics should satisfy the following requirements [1]:

- **Universality**: each individual should have the selected characteristics.
- **Distinctiveness**: the characteristic should facilitate discrimination among different individuals.
- **Permanence**: the characteristic should be sufficiently stable, with little or no change over time.
- **Collectability**: the characteristic should be capable of being captured and quantised.
- **Performance**: measurement of the characteristic should achieve an acceptable level of accuracy and should take account of other factors such as cost and operational speed.
- **Acceptability**: individuals should be willing to permit use of the characteristic and to present it to the system.
- **Circumvention**: the characteristic should be sufficiently robust to fraudulent approaches.

In contemporary society, there is a need for a reliable real-time system for identity recognition in many contexts, including computer network access, border control, and financial transactions. In improving security in these and other settings, the most commonly used biometric modalities are face recognition, fingerprinting, and iris recognition. Any reliable biometric recognition system usually consists of two main stages: (i) enrolment and (ii) recognition. At the enrolment stage, data are captured from individuals, and the extracted features are stored as gallery templates. At the recognition stage, the system establishes the identity of individuals by comparing the similarity between the gallery (template) and probe.

Depending on the application scenario, recognition systems can operate in either of two modes: *verification* or *identification*. In verification mode, the system attempts to validate an individual's claimed identity by comparing the captured image with the image in the database (template)—in other words, the system performs a *one-to-one*

match. The claimed identity is verified using a pre-set threshold that determines the system's decision, as for instance in mobile applications for access control. In contrast, in identification mode, the system seeks to determine an individual's identity by comparing the probe capture against all gallery templates stored in the database. The system produces a set of similarity scores for matching, and the probe capture is assigned to the database item with the highest similarity score. In this case, the system performs a *one-to-many* match, also referred as 1: N , where N is the number of samples in the gallery template. This mode is commonly used in forensic or surveillance applications [2].

1.2 Skin Texture as a Biometrics

Biometric traits including fingerprints, irises, and faces have been widely used in practical applications that range from phone access to border control. Face recognition systems are also a proven component of biometric technology. The key advantages of face recognition systems are their non-intrusiveness and the fact that they do not require the test subject's cooperation because of their ability to identify the individual in a crowd. However, despite the successes of face recognition, some challenges remain to be addressed. For example, in typical applications such as the identification of individuals in video surveillance frames or in images captured by handheld devices such as mobile phones, where without user cooperation, a face may be only partially captured. In such scenarios, face recognition systems will probably fail to identify individuals, as this requires the whole face image. One possible solution is to exploit information from the available part of the face image, and in such these scenarios, skin texture features are a key source of significant biometric information.

Skin texture has been shown to contain plentiful and detailed information that can be used in various contexts, including health status [3] and identity information [4] [5] [6] [7]. In medical research, skin texture has been studied in many contexts, including computer-aided diagnosis in dermatology, where the patient's skin texture can be used for diagnostic purposes [3]. In this latter diagnostic application, skin texture can be used for biomedical evaluation of skin treatments and their effectiveness. With the rapid developments over the last decade, well-known biometric technologies such as

fingerprints [8] and palm prints [9] have employed skin texture to detect identity, using fingerprint skin to extract texture features for a person identification [10]. More recently, skin texture has been used to help identify the pattern of blood vessels hidden in colour images [11].

It is generally believed that facial skin appearance is influenced by factors such as ageing [12], variations in illumination [13], facial expression [14], makeup and gender [15]. Skin appearance also changes noticeably over the lifespan [16]. In general, young people have smoother skin than older people, and ageing is clearly revealed by wrinkles or other facial marks. Environmental and social factors such as lifestyle, weather changes, smoking, and stress all cause major variations in the appearance of human skin.

Forensic demands and the weaknesses of other biometric traits make it essential to develop recognition systems based exclusively on skin information, and a few studies have investigated the use of human skin texture information to identify individuals. The earliest work proposed by Lin et al. [5] investigated the possibility of using skin texture information as additional source of distinctiveness to achieve performance improvement for face recognition systems. Skin marks (e.g. scars, moles, freckles) have also been studied as a means of improving the overall accuracy of face recognition systems [17].

In general, a skin biometric system commonly consists of the following components: image acquisition, skin region localisation, skin region segmentation, skin feature extraction, and skin classification. Figure 1.1 describes a generic skin biometric system.



Figure 1.1 Schematic diagram of a skin biometric system

In the **image acquisition** step, the system captures the skin image and quantises the sample into digital form for further processing.

Skin region localisation determines the region of interest (ROI), using facial landmarks such as the outer corner of each eye, the nasal bridge, and the corners of the mouth. Automatic localisation is often performed by identifying a set of geometric facial

measurements, using methods such as Active Shape Models [18], Active Appearance Models [19], or Chehra Face Tracker [20] to detect facial landmarks.

Skin region segmentation crops the localised ROI. To validate the usability of the ROI, it is submitted to skin purity assessment scheme. Skin colour pixel detection techniques are usually applied to categorise image pixels as skin or non-skin on the basis of quantised pixel colour. This step is sometimes applied prior the skin purity assessment scheme, ensuring that the acquired skin data can be reliably processed for further steps.

Feature extraction constructs features from pure skin regions. Extracting and generating a compact and expressive representation (known as a feature set) plays a vital role in skin biometric systems. The feature set is usually stored as a template in the system database for further comparisons.

Skin classification compares features extracted from different approaches to generate matching scores for biometric recognition. As noted earlier, the objective of identification mode is to determine the user's identity among a set of individuals, in which the claimed identity is compared against the entire template in the database (one-to-many matching). In verification mode, the claimed identity is compared only to the template corresponding to the claimed identity (one-to-one matching).

1.3 Motivation and Objectives of the Research

The present study aims to develop the usability and practicality of facial skin biometric technology for wider application in forensic, surveillance, and security contexts where a system must be capable of identifying individuals in crowded scenes. In such scenarios, one critical issue is that traditional biometric traits (e.g. fingerprint, iris, face) are either not always available and/or cannot be captured. Face recognition systems may help to mitigate such issues, but in such scenarios, the face image may be partly occluded by objects or accessories (sunglasses, hat, scarf, etc.), or by facial hair. As a composition of micro-patterns, skin information is a potentially useful source of biometric information for person recognition in such cases. Additionally, the development of multimedia hardware such as HDTV and high-resolution digital cameras means that an image of the

skin can readily be obtained, making it easy to extract and analyse subtle and detailed information. It has been suggested that relatively high-resolution images are needed to capture the distinctiveness of skin texture. For example, in [21] and [7], skin features extracted from high-resolution images for person verification yielded promising results. Conversely, the use of lower resolution images can significantly undermine skin biometric performance, and overcoming noise degradation to improve skin system performance remains an open problem. A further issue is that matching of skin information captured at different resolutions has yet to be investigated.

On that basis, the present thesis addresses the design of a robust framework to improve automatic skin-based biometric person recognition systems. The thesis also explores the effect of skin resolution changes on recognition performance. The key objectives of the study are as follows:

- to design an automated skin purity assessment scheme prior to extraction of skin information for analysis of biometric features
- to develop an effective skin biometric system for different facial regions
- to explore which facial skin regions provide the best biometric information
- to investigate the system's performance using skin features at different resolutions
- to establish an adaptive skin-based biometric framework to extract only skin image pixels from a noisy region of interest
- to explore the possibility of using only a small part of the region of interest at arbitrary resolution for a person identification

1.4 Contributions

This thesis is concerned with the development of a framework for exploiting skin texture features for person recognition. The main contributions are summarised as follows:

First, a skin pixel detection technique is designed and implemented to assess the potential of the skin region of interest for biometric processing. This step removes the need for human interaction and enhances the accuracy of the process. The technique is based on skin colour thresholding at pixel level to detect distortions such as hair or other artefacts. Two colour spaces were adopted and combined with other techniques in the literature. It is shown that the proposed skin detection's performance returned comparable results (Chapter 4).

Second, a fully automated localisation technique of facial skin regions including Forehead, Right cheek, Left cheek, Chin is presented. Geometric measurements for facial skin region of interest are proposed to accurately determined skin regions within the face image (Chapter 5).

Third, an automated skin-based biometric system is presented to investigate the use for texture information of each facial skin region, evaluated, and compared with other published results. Skin texture features extracted using different methods were explored and compared. The experimental results demonstrate that the proposed system achieved an improved performance and reliable results. Skin regions which have a more closely planar surface (e.g. forehead, and chin region) provided better recognition than less-planar regions (e.g. cheek) (Chapter 5).

Fourth, a novel scheme for using facial skin texture features at different resolutions is proposed and implemented. Skin features extracted from images of different resolutions are investigated to identify a person from different scales. Information fusion at feature extraction and score levels are analysed. Experimental results show that not only the high-resolution skin image can provide valuable biometric information, but also low-resolution skin image can contribute significant biometric information (Chapter 6).

Fifth, an adaptive skin-based biometric technique to exploit of the available skin information from the region of interest is proposed. The feasibility of the proposed technique is adapted and implemented for evaluating arbitrariness of both skin region location and skin resolution for personal identification (Chapter 7).

1.5 Thesis Structure

As shown in Figure 1.2, the thesis is organised in eight chapters, which can be briefly summarised as follows.

- **Chapter 2** includes a comprehensive survey of the literature on face recognition systems and the current challenges posed by skin texture features. Historical developments in face recognition systems are also discussed. Human skin and related issues covering the state of the art are presented. Finally, frequently used databases for the evaluation skin biometric systems are reviewed.
- **Chapter 3** presents a comprehensive, up-to-date review of advances in texture analysis over the last two decades. The chapter covers different aspects of texture analysis, including problem descriptions, texture feature extraction methods, and texture feature classifications. The measures commonly used for assessing the performance of biometric systems are also presented.
- **Chapter 4** describes the techniques used for face skin detection. The chapter includes a comprehensive, up-to-date review of previous work on skin pixel detection techniques. Skin modelling schemes for purity assessment and the database used for system validation are described in detail. The proposed technique's performance is also discussed.
- **Chapter 5** presents the design and implementation of a skin-based biometric system. The system is established to explore the potential of facial skin regions for person identification. Skin pixel detection and purity assessment for skin images that support skin biometric processing are also discussed, along with performance evaluation techniques. The chapter then presents an advanced purity assessment scheme of facial skin regions to further improve the performance of the skin texture-based biometric recognition system. Experimental results suggest that skin biometric performance is better for the forehead region compared to other facial skin regions.

- **Chapter 6** reports the design, evaluation, and validation of the skin biometric system at different resolutions. Approaches used to scale skin face images using interpolation methods to generate different skin resolutions are described in detail. Experimental setup and experimental results are also discussed, along with the influence of skin image resolution on person recognition.
- **Chapter 7** describes the technique used to process available information from a part of the skin image, where the other part is inadequate for biometric purposes. The chapter also details extensive experiments on the proposed adaptive skin biometric technique. Finally, the ability of matching different skin resolution images at arbitrary locations within the ROI is investigated, and the results of the experiments are presented and discussed.
- **Chapter 8** draws conclusions from this research and makes recommendations for future work.

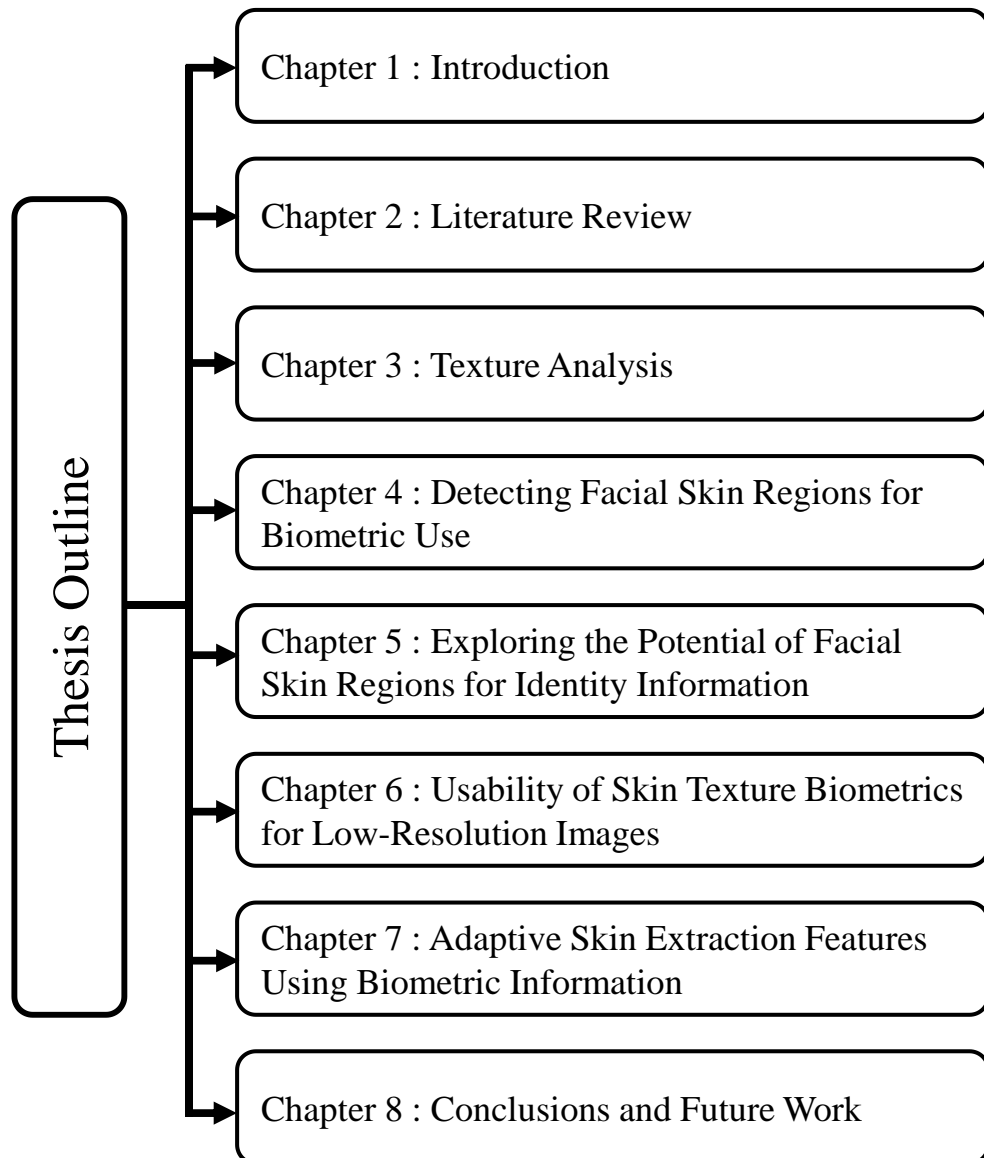


Figure 1.2 The outline of the thesis

Chapter 2

Literature Review

2.1 Introduction

This chapter revisits previous works on facial skin-based biometric systems for identity information, beginning with a brief but precise description of challenges for face recognition systems and the historical development of these systems. As the thesis focuses on building and developing an algorithm for automatic facial skin-based biometrics, much of the chapter is concerned with the nature of human skin and related issues. Generic facial skin systems are also detailed, and the chapter ends with an outline of the databases most commonly used for skin-based recognition.

2.2 Challenges for Face Recognition

Face recognition is now among the most widely used modalities in biometrics. While the accuracy of face recognition systems has improved significantly over the last two decades, some limitations and technical challenges remain. A state-of-the-art review highlights the following challenges for face recognition systems:

- Unconstrained environment
- Real-time recognition

- Ageing
- Changes in face position, size, and pose
- Changes in facial expressions
- Low-resolution facial images due to camera quality, lens characteristics or distance between camera and object
- Accessories such as glasses, caps, and hair
- Partial occlusion of the face
- Gender classification
- Facial changes at a distance

Facial skin texture provides additional information that can help to overcome some of these challenges. Facial skin can be represented by structures on its surface such as wrinkles, scars, pores, birthmarks and moles. On that basis, this chapter reviews the most recent research on face recognition systems, focusing on the use of skin information for recognition of individuals. Following a general introduction to historical development in face recognition systems, the focus then narrows to progress in facial skin biometric systems as reported in the literature.

2.3 Historical Developments in Face Recognition Systems

Face recognition systems have been widely used since the 1960s, and progress has been made in applying them in real-world settings [2]. Algorithms have developed rapidly, and the increasing number of available databases has contributed to system improvements. Another major factor in these improvements is the greater range and availability of imaging devices [22]. Image sensors have become smaller and less costly, making it possible to embed them in many electronic devices. Figure 2.1 illustrates some fundamental milestones in the progress of face recognition systems and camera sensors over time, including the evolution of face recognition algorithms and turning points in the development of face acquisition systems.

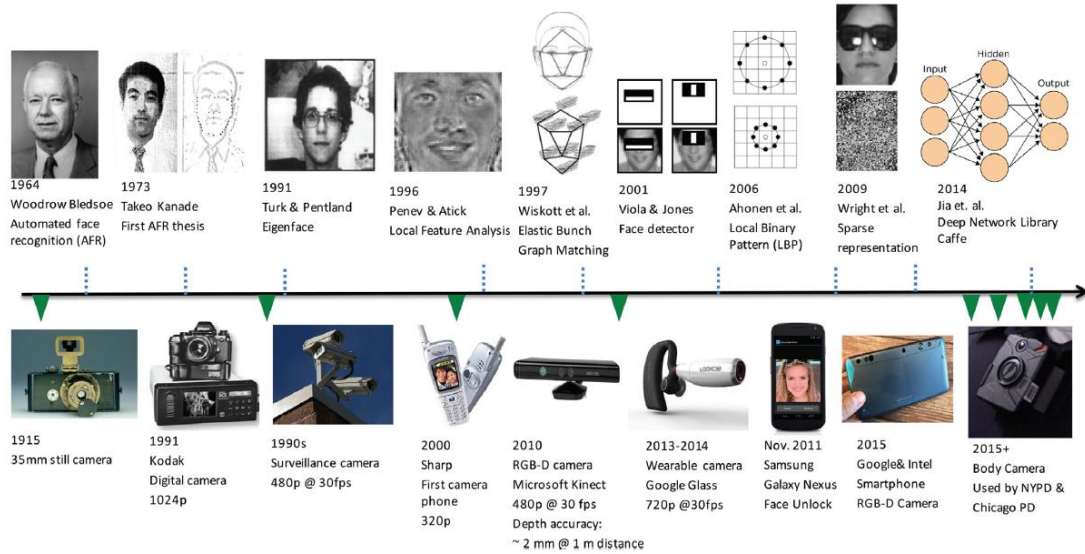


Figure 2.1 Some significant milestones in the history of face recognition systems [2]

2.4 Human Skin

Skin can be defined as the outer tissue of the body. Humans usually take great care of their skin's appearance, and this is of great interest in several fields of technology and science, including medicine, cosmetology, computer graphics and computer vision (see Figure 2.2). While briefly addressing some of these areas, this section is principally concerned with developments in computer vision.

Skin texture images can be corrupted or changed as a result of factors such as hair, cosmetics, spectacles, acne, freckles, razor burns or changes of lighting. Skin texture is also affected by ambient environment (e.g. sunlight causing sunburn, skin peeling, darkening freckles etc.), as well as by diseases or treatments (e.g. skin allergy, eczema, vitiligo, drug eruptions, chicken pox). Biometric studies have devoted great attention to skin appearance as a means of identifying individuals, as for instance in the extensive research on fingerprint recognition systems. Skin on other parts of the body has also been used for personal identification, including the palm, leg, arm and thigh [23].

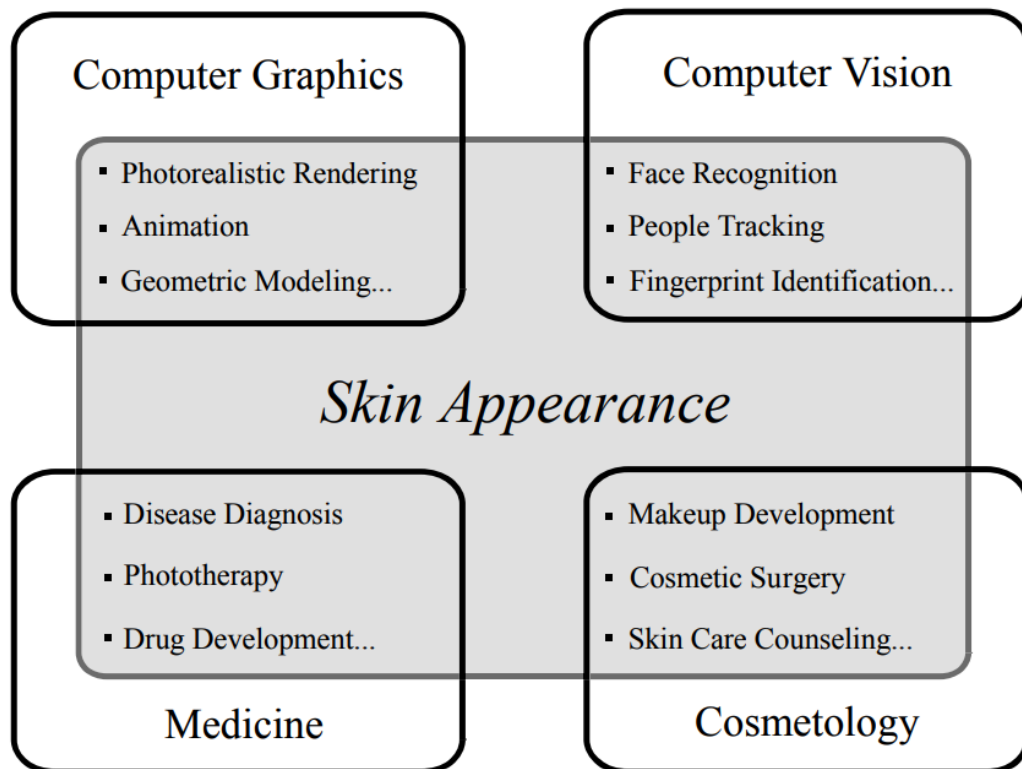


Figure 2.2 Fields of science and technology utilising skin appearance [24]

2.4.1 Skin Texture and Disease Detection

Skin texture has been used in medical image analysis to detect a wide range of diseases. For example, breast cancer can be detected at an early stage through classification of parenchymal density and detection of micro-calcifications [25]. In dermatology, skin texture information has been used in developing a method of bidirectional imaging to analyse the skin's surface [3]. Five dermatological disorders have been classified on the basis of skin texture: acne, congenital melanocytic nevus (medium-sized), keratosis pilaris, psoriasis and acute allergic contact dermatitis [3].

2.4.2 Skin Texture and Cosmetology

Females are more likely to use makeup in their daily life to remove facial flaws or to change the appearance of their face for cosmetic purposes. Additionally, makeup is still widely used as a socially acceptable and cost-efficient option. However, makeup can create problems for face recognition systems, as it may change the appearance of the individual's skin texture, and researchers have recently begun to address this issue. For

example, Guo et al. [26] used four features to study the relationship between faces with makeup and those without: skin colour tone, smoothness, texture and highlights. Using a correlation-based scheme, they explored makeup-invariant face authentication using two images of the same person: one with makeup (of different kinds) and one without. The results indicated that a combination of colour and smoothness features in the image helped to detect the presence of makeup.

Dantcheva et al. [27] tested the effect of facial makeup on an automated biometric system and found that it significantly changed the facial appearance by altering colour, contrast and texture, undermining the performance of the face recognition system. In a study of facial aesthetics based on soft biometrics and photo quality, Dantcheva and Dugelay [28] expanded the range of face recognition to include artificial features such as makeup and hairstyle. They claimed that their method can detect the modification of aesthetics in facial images. Chen et al. [29] studied the binary classification of faces with and without makeup, including shape, texture and colour features as a means of detecting facial makeup from an unconstrained facial image. Two unconstrained female datasets (the YouTube makeup database (YMU) [27] and the Makeup in the Wild database (MIW) [29]) were used to evaluate their proposed system. The experiments investigated the effects of facial pose, illumination, expression and image resolution, using Gabor wavelets and local binary pattern (LBP) texture descriptors to characterise micro-patterns and micro-structures in the facial images. Support Vector Machine (SVM) and AdaBoost classifiers were employed for classification. Recent work by Batool and Chellappa [30] reported a technique for detecting wrinkles and other imperfections in the facial skin, based on texture-orientation field and Gabor filter features. The algorithm detected wrinkles and inpainted irregularly shaped gaps to remove them.

2.4.3 Skin Information for Personal Recognition

Skin biometric systems (i.e. approaches that typically use only skin information for personal recognition) are sometimes embedded in other systems to enhance the overall recognition performance. For example, Klare et al. [31] [32] divided facial features into three main levels for representing face images: 1) global appearance, which can be

assessed using Eigenfaces; 2) geometric and structural measurements to differentiate between two similar faces (often extracted by SIFT, LBP descriptors, and/or Gabor wavelets); and 3) skin irregularities and micro-features (e.g. scars, facial marks) that may be useful in cases of inter-class similarity (e.g. identical twins) or partial view (e.g. occlusions). The authors examined the uniqueness of level (2) using Multiscale Local Binary Patterns (MLBP) and Scale-Invariant Feature Transform (SIFT) descriptors, and of level (3) using facial marks to assess inter-class similarity. Their system was developed to analyse the uniqueness of several facial components (e.g. eyes, eyebrows, nose, mouth), using linear discriminant analysis (LDA) to address two questions in particular: (i) whether facial feature subspaces can be learned in order to distinguish between twin pairs and (ii) whether these subspaces affect face recognition performance in non-twin comparisons.

Partial faces are encountered in many real-world scenarios, especially in unconstrained image capture environments involving surveillance cameras or handheld devices (e.g. mobile phones). As only a small part of the person's facial image may be available at the crime scene, that part must be fully exploited in order to identify the individual concerned. In [33], the authors claimed to have developed an approach to partial face recognition that could distinguish an arbitrary partial face without face alignment. The approach used a variable-size description to represent each face by a set of keypoint descriptors, based on holistic or partial face images represented by a large dictionary of gallery descriptors.

Su et al. [34] proposed an approach that combined global and local facial features as a hierarchical ensemble classifier (HEC). At each position of the facial image, the Gabor wavelet transform was exploited for local feature extraction. These features were then assembled into a number of so-called local Gabor feature vectors (LGFV), each corresponding to a local patch of the facial image.

Automated facial skin biometric techniques address the need for forensic, surveillance and security applications in crowded scenes. In particular, it has been reported that two persons (even twins) may be unlikely to share the same skin characteristics [35]. Several studies have investigated the possibility of utilising skin texture as a source of information for identifying individuals. For instance, Pierrard et al. [6] proposed a

method for extracting local skin irregularities such as moles and birthmarks from facial images, which would then be analysed to determine the individual's identity. However, a major limitation of that framework was that human interaction (i.e., manual detection) was required to define facial landmarks (e.g. nose, corners of both eyes). Lin and Tang [5] proposed a multilayer system for face recognition, in which facial skin features constituted one of these layers. They defined skin texture as repetitive texture units, and a Gabor filter was used for feature extraction and skin texture representation. Linear Discriminant Analysis (LDA) was applied to reduce the irrelevant variations. Their results suggested that skin texture can be used as an auxiliary feature. Park and Jain [17] also proposed a person identification technique based on facial marks such as scars, moles, and freckles. The active appearance model (AAM) was used to localise primary facial features such as eyes, eyebrows, nose, mouth, face boundary. Laplacian of Gaussian (LoG) operator was applied to detect facial marks. However, as not all users exhibit such facial marks; these are insufficient to establish the identity of all individuals in a real-world application.

Using high-resolution facial images, Dong et al. [7] recently proposed a pose-invariant method of face verification that is robust to alignment errors and pose variations. In their proposed system, eye and mouth regions were used to locate facial landmarks. These landmarks were manually demarcated. The feature extraction step involved Principal Component Analysis (PCA) and the Scale Invariant Feature Transform (SIFT) descriptor [36] (PCASIFT); they named their method Pore-Principal Component Analysis (PCA)-Scale Invariant Feature Transform (SIFT) (PPCASIFT). In [21], they proposed another descriptor called Pore-SIFT (PSIFT), using skin from the hairless cheek region of the facial image to evaluate the algorithm. The algorithm's performance was evaluated using three public databases: the Bosphorus dataset (1400×1200 pixels) [37], the Multi-PIE dataset [38] (face images of this dataset with a minimum resolution of 600×700 pixels provided sufficient detail to reliably match two skin images from the same subject), and the FRGC v2.0 dataset [39], all comprising high-resolution images. The method was compared with the Eigenface method (PCA), the Gabor feature with PCA (Gabor+PCA) method, the LBP method and the LBP feature with PCA (LBP+PCA) method.

Recent advances in imaging technologies and the wide availability of digital cameras have made many more human images available for police and forensic investigations. Different parts of human body can be easily captured at the crime scene. In such cases, skin texture may be the only option. For example, several studies have investigated the use of blood vessel patterns and skin marks [40] [41] for personal identification. Although blood vessel patterns can be captured using near infrared imaging systems, the visibility of these patterns depends on factors such as the thickness of the subcutaneous fat layer and the skin's pigmentation level. The other limitation is that while skin marks are easier to observe, they are likely to be covered by hair or other items. Another study [23] used various parts of the human body, including inner forearms and inner thighs, to extract distinctive skin information for identification purposes. The goal of their study was to develop a skin texture system for forensic applications. Local binary pattern (LBP) and Gabor filters were used to extract skin texture features, and partial least squares regression (PLS regression) was used to classify skin texture.

Skin texture on the back of the hand has also been investigated for the purposes of identification and gender classification [42]. In that study, hand images were collected using a high-resolution camera, yielding a high quality skin texture image of 450 dpi (dots per inch).

Moreover, skin texture features are exploited in different aspect of human recognition. For example, Jizheng et al. [43] proposed an algorithm for facial expression recognition using structural characteristics and texture information obtained from the face image. Using the active appearance model (AAM) to determine facial landmarks, three facial features—skin deformation energy parameter, connection angle ratio coefficient and feature point distance ratio coefficient—were used to differentiate between individuals. Facial expressions were then classified by exploiting the radial basis function neural network.

Table 2.1 shows performances of the most relevant work in the literature, including both identification and verification biometric evaluation systems. The uses of skin images vary according to the application, but the literature reveals that more work remains to be done on face skin images.

Table 2.1 State of the art n facial skin texture biometric performance

Publication		Face representation	Database (No. of subjects)	Accuracy
Authors	Year			
Pamudurthy et al. [44]	2005	Skin textures (pores) using skin correlation (DISC)	Private	NA
Lin and Tang [5]	2006	Skin textures using SIFT and R-LDA	XM2VTS (295)	Recognition rate = 0.669
Pierrard and Vetter [6]	2007	Irregularities in facial skin (e.g. moles, birthmarks)	FERET (194)	Recognition rate > 87%
Yi et al. [45]	2009	Skin texture using LoG and binary features	MBGC (114)	Verification rate = 77% at FAR = 0.1
Jain and Park [46]	2009	Facial marks (e.g. freckles, moles, scars) using LoG and morphological operators	FERET (426 images, 213 subjects); Mugshot (1,225 images, 671 subjects)	Recognition rate = 93.90% (FERET) 93.14 % (Mugshot)
Park and Jain [17]	2010	Global and local texture using AAM, LoG, PCA	FERET (10 213)	EER= 3.853%
Klare et al. [32]	2011	Facial marks using SIFT and MLBP	Twin data (174)	TAR = 50% at FAR = 1.0
Al-Qarni and Deravi [47]	2012	Skin texture using Gabor and LBP operators	XM2VTS (84)	Recognition rate = 68.16%
Li et al. [7] [21]	2015	Pore-scale skin features using PSIFT and PPCASIFT	FRGC v2.0 dataset	EER = 5.51% (PSIFT) and 7.36% (PPCASIFT)
Weng et al. [48]	2016	Skin texture using RPSM	PubFig (140)	Verification rate = 66%

2.4.4 Skin texture biometrics in Low-Resolution Images

There is clear evidence that image resolution is important for the performance of biometric recognition systems. For example, in identifying a person by means of surveillance cameras or cell phone face recognition, challenges that may affect system performance include expression, age, illumination, pose, occlusion, and distance, and these have not been fully resolved [49] [50].

Image resolution is a critical factor in recognition performance. In commercial face recognition systems, facial images are usually considered high-resolution when the distance between the eyes (the *inter-pupillary* distance) is greater than 60 pixels. However, government systems (e.g. passport or personal identification technologies) require an inter-pupillary distance of at least 90 pixels for successful recognition [51]. As the best practice, the ISO/IEC JTC 1/SC 37 committee recommended that this distance should be 120 pixels, which would require 12 pixels per millimetre (300 pixels per inch). On that basis, an *inter-pupillary* distance of less than 60 pixels would be considered low-resolution.

Matching two face images of the same subject at different resolutions (e.g. different distances) is of particular interest in forensic applications, and some work has been done on recognising faces in high-resolution (HR) and low-resolution (LR) images. LR images captured by surveillance cameras [52] [53] are often indistinct, making individuals difficult to be recognised. Using traditional methods [54] [55], HR images usually provide more detail than LR. To exploit LR images, computer vision and machine learning researchers have employed a technique known as super-resolution (SR) to generate HR images from LR [56]. There are three categories of SR: learning-based, functional-interpolation, and reconstruction-based. All of these methods work only with the whole face image [57].

Earlier work [58] achieved face recognition by applying SR to low-resolution images. However, accuracy decreased dramatically when whole face image resolution was less than 16 x 16 pixels in size. It has been reported that some algorithms require a resolution of at least 32 x 32 pixels [59] to recognise individual faces, but recognition may be possible below this resolution in real-world applications [59] [60]. Early

research addressed this problem by introducing different SR algorithms [61] [62], but to date, low-resolution (e.g. captured from a distance) partial face images have not been considered. Instead, most previous studies have focused on whole face images, which are often unavailable. The partial view of the face at a distance has not been investigated because the whole face image has not delivered acceptable performance at a distance.

Today, high quality facial images can be easily captured from different distances, using imaging devices such as high-resolution cameras, phones, webcams, and other inexpensive computing resources. For example, pan-tilt-zoom (PTZ) cameras can capture a high-resolution facial image from up to 12 meters, with automatic tracking and close-ups [63]. It also has been reported [59] that face recognition systems are developing rapidly, with increasing access to high-resolution equipment [59]. All of these factors facilitated this exploration of the use of a small part of the facial image (such as facial skin) to identify individuals at a non-fixed location/distance.

HR digital cameras have facilitated more sophisticated analysis of skin texture features, and extensive research has confirmed that skin information extracted from a small part of the facial image can provide meaningful biometric information [64] [65]. It is opportune, then, to investigate the possibility of matching the skin information embedded in high- and low-resolution images. The present work investigated the use of skin texture features at different resolutions to identify an individual, assessing how the proposed framework performs in difficult conditions (e.g. using low-resolution images) and how image resolution affects recognition. The study contributes to the enhancement of skin-based biometric systems, and to the best of author's knowledge, this is the first work to clarify the relationship between resolution and use of facial skin images for person recognition.

2.5 Generic Face Skin Biometric System

As skin texture is one of the most important sources of information about the face, it is important to understand the strategies used to obtain this information. In general, every biometric system initially follows the steps in Figure 1.1. Most face images are captured in the visible light spectrum, and sensors are likely to differ across imaging devices, causing changes in the appearance of facial skin. Face images can be categorised into

two main types of feature [66]. The first category refers to global features that describe the overall shape of the human face, including eyes, nose and mouth, which requires the whole face image for classification purposes. The second category refers to local features that describe the structure of face image components, each of which can be used independently to classify individuals. Using these components independently is advantageous in addressing some of the challenges of face recognition, such as occlusion, pose and facial expression.

2.5.1 Face Detection

The objective of face detection systems is to detect and locate faces in an image. Any face detection system should be able to deal effectively with arbitrary variations such as pose and occlusion. There is a broad agreement that the Viola and Jones' face detector is by far the most successful face detection algorithm [67]. This algorithm is based on Haar-like features and a cascade AdaBoost classifier [68]. Another typical approach uses colour images and skin-based detection. Greyscale facial images can be also detected using knowledge-based, feature-based, template-based and appearance-based methods. For more details of these methods, the interested reader should refer to [69] [70] [71].

Recently, Liao et al. [72] proposed a method to identify face images using unconstrained face detection with arbitrary facial variations. The technique is based on a pixel-level feature called Normalized Pixel Difference (NPD), which is computed as the ratio of the difference between any two pixel intensity values to the sum of their values. It was claimed that this face detector can handle issues such as pose variation, occlusion and lower image resolution in unconstrained scenarios [72].

2.5.2 Face Normalisation

The preliminary processes of facial feature alignment include face normalisation, which is often implemented once the face image is detected. This technique can be applied to enhance face recognition systems, where normalisation can be either geometric or photometric. Geometric normalisation is usually applied to reduce the rotation of face images (head pose etc.) while photometric normalisation is applied to eliminate the

effect of illumination. These representative methods to normalise face image are reviewed as follows.

2.5.2.1 Geometric Normalisation

Normalisation of facial images can be used to alleviate the influence of head pose when the region of interest (ROI) is cropped. This technique is applied to the eyeballs to fix the tilt of the facial image. Formally, let (x_r, y_r) be the coordinates of the right eyeball centre and (x_l, y_l) the coordinates of the left eyeball centre. The image can be rotated clockwise or counter clockwise. The direction is determined by using the polarity of difference of vertical distances $(y_r - y_l)$ between the coordinates of the two eyeball centres. The new coordinates of the eyeball centres are computed as follows:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (2.1)$$

where θ is the angle of image rotation, calculated as:

$$\theta = \tan^{-1} \left(\frac{y_r - y_l}{x_r - x_l} \right) \quad (2.2)$$

2.5.2.2 Photometric Normalisation

Photometric normalisation is used to eliminate any illumination effects, often after geometric normalisation. The technique can take two forms, the first of which is to train face images to learn a global model of possible illumination—for instance, by linear subspace analysis [73]. The second type uses conventional image processing transformations that require no training sample, such as the Retinex [74], Histogram Equalisation, pre-processing sequence [75] and Limited Adaptive Histogram Equalisation approaches.

2.5.3 Feature Extraction and Classification

As this thesis focuses on facial skin texture information, feature extraction will be described in Chapter 3 as an element of skin texture analysis. Classifiers used to classify skin texture information will be also described at that point (see Sections 3.4 and 3.5).

2.6 Databases

There are several facial skin image datasets, which are derived from face image databases. Each of these was collected to deal with specific challenges in human identification. The most widely used facial image databases are listed in Table 2.2.

Other human body skin databases available to the research community are also collected for either forensic or crime-related applications. Among recent datasets of this kind, the NTU datasets [76] include the NTU Internet Image Set v1 [76], the NTU Human Back Skin Dataset, the NTU Human Chest Skin Dataset, the NTU Human Inner Forearm Skin Dataset v.2, the NTU Human Inner Thigh Skin Dataset v.2, [77] [78], the NTU Lower Leg Skin Dataset [79] and the NTU Tattoo Dataset [80]. All of these datasets were collected from prisoners, gangsters or offenders to enable law enforcement agencies to investigate suspects and to identify criminals.

In addition, NIR images were collected using an NIR camera [23]. For example, Images in NTU colour forearm dataset have corresponding NIR images in the NTU NIR dataset. Datasets were collected from different subjects who were mainly Chinese, Malay, Indian, Vietnamese, and Caucasian in Singapore. The interested readers are referred to [23] [76] [77] [78] [79] [80].

2.7 Summary

To exploit facial skin information, challenges for face recognition systems and their development over the last decade must first be understood by undertaking a comprehensive literature review. This extensive overview of facial skin biometric addressed topics of general relevance, and a generic face skin biometric system was briefly but precisely discussed.

Table 2.2 Widely used face image databases

Database	No. of subjects	No. of images per subject	No. of images in the database
XM2VTS	295	4	1180
FERET	1,199	≈ 12	14,126
AR	126	≈ 32	4,000
Oulu Physics	125	16	2,000
AT&T (formerly ORL)	40	10	400
FRGC	625	80	50,000
Yale	15	11	165
Yale B	10	576	5,760
SCface	130	32	4,160
SCface B	130	31	4,160
PIE	337	-	750,000
Cohn-Kanade AU-Coded Facial Expression	100	-	500
MIT-CBCL	10	200	2,000
JAFFE	≈ 30	7	213
BioID Face	23	66	1,521
Caltech 10000 Web Faces	$\approx 10,000$	-	10,000
CAS-PEAL face	1,040	-	99,594
Georgia Tech Face Database (GTFD)	50	15	750 (no longer existing)
Indian Face Database	40	-	more than 440
Face Detection Data Set and Benchmark	2,845	-	5,171
PIE Database, CMU	68	-	41,368
UCD Colour Face Image Database for Face Detection	≈ 299	1	299
UMIST Face Database	20	-	564

The most widely used databases for evaluating skin biometric systems were also presented. The next chapter describes skin texture analysis and its use in the field of biometrics.

Chapter 3

Previous Work on Texture Analysis

3.1 Introduction

This chapter presents a comprehensive literature review exploring previous works on texture analysis. The state of the art review includes all relevant references and material related to challenges, historical developments and approaches to feature extraction and classification. Because the present research focuses on local feature-based methods, the chapter only emphasises such approaches. As well as providing useful and necessary background knowledge that informs the present work, the literature search helped to ensure that this research programme will make a worthwhile contribution.

3.2 History of Texture Analysis

Since the 1960s, texture analysis has been widely studied as a fundamental and challenging problem in many fields, including computer vision and pattern recognition [81]. The need to understand how texture perception works in human vision and its use in a wide variety of applications has prompted intensive research. By the 1990s, human vision research had broadened to include problems such as texture classification, segmentation, synthesis and shape. Texture analysis has been applied in areas such as medical image analysis [82], analysis of satellite images [83], face analysis [84] [85]

and biometrics [86] [87]. In the 1990s, Ojala et al. [88] conducted a comprehensive evaluation of texture feature performance, resulting in a texture feature descriptor called Local Binary Pattern (LBP) [88], which represents a major milestone in the history of texture analysis. Figure 3.1 overviews progress and major milestones in the field of texture analysis.

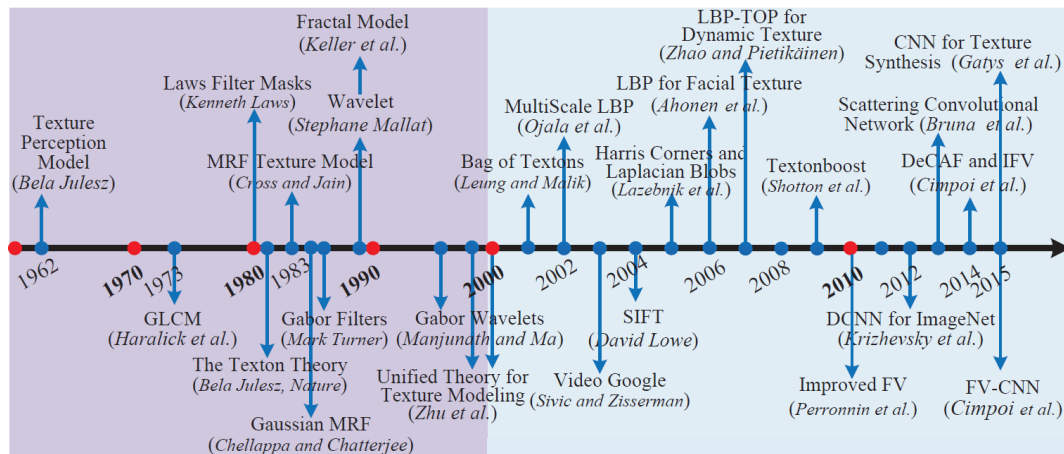


Figure 3.1 Timeline of work on texture representation [89]

3.3 Review of Texture Analysis Definitions and Challenges

3.3.1 Texture Analysis

There is no single definition of the term ‘texture’ on which all scientific and academic researchers agree, presenting a challenge for many researchers in the fields of image processing and computer vision. This sub-section briefly reviews some existing definitions.

- *Texture* can be understood as repetition of a greyscale or colour pattern on an image [90].
- *Texture* is defined in terms of biological vision as an important cue that helps humans to distinguish between objects. [91].
- Visual *texture* refers to an image surface containing variations of intensity that form certain repeated patterns [92].
- Image *texture* is a rich source of visual information about the nature or shape of physical objects [93].

- *Textures* can be defined as complex visual patterns composed of both entities and sub-patterns that contain a degree of characteristic brightness, colour, slope and size [94].
- *Texture* constitutes a macroscopic region; its structure can be attributed to repetitive patterns of elements or primitives organised according to a placement rule [95].
- *Texture* can be understood as the appearance, structure and arrangement of the parts of an object within an image [96].
- *Texture* can be defined as an attribute of a field that has no components that appear enumerable. Overall properties of texture include coarseness, bumpiness and fineness [97].

3.3.2 Issues Related to Texture Analysis

The various approaches to texture analysis and feature extraction usually proposed to address five common categories of problem:

1. *Texture segmentation* is the technique of separating an image into several regions. Its many applications include interpretation of satellite images, traffic control systems, diagnostic medical imaging, measuring tissue volumes and studying anatomical structure. The technique can be divided into two distinct approaches: region-based and edge-based. While region-based texture segmentation uses textural similarity between neighbouring regions, the edge-based method detects boundaries with adjacent regions that are dissimilar in texture. The latter approach does not therefore need to identify the number of texture regions in an entire image [98, 99].

2. *Texture classification* is used to assign unknown objects to one of the known texture classes. This can be implemented in real-world applications to allow subjects to be viewed as a sort of texture, as for instance in fabric classification [100], rock classification [101] and identification of wood species [102]. The process involves two steps: a training phase and a test phase. In the training phase, a set of grey-level images is derived from every texture class, and their properties are captured by means of the selected texture analysis approach, yielding a set of textural features for every image to build the model of each texture class. In the testing phase, classification algorithms are

used to compare the feature vector with all vectors obtained in the training phase in order to assign the test sample to the class that returns the best match.

3. *Texture defect detection* is a method of determining the location of pixels exhibiting a notable deviation in intensity values in a textured image. This has been applied in a number of ways to the problem of inspection of different surfaces, including wood, steel, paper and textiles, attracting great interest from industry as a means of replacing the repetitive process of manual inspection [103, 104].

4. *Texture synthesis* provides descriptions for modelling image texture; the model can then be used to represent the texture, usually in computer graphics [105]. In [106, 107], texture synthesis based on the first-order and second-order properties of joint wavelet coefficients yielded good results, but high frequencies were not reproduced in some highly structured patterns.

5. *Shape from texture* shows image properties that allow detection of the shape of an object in 3D—that is, reconstruction of 3D surface geometry from texture information such as intensity differences on the object’s surface, orientation of edges and corners and shadowing effects [108-110].

3.4 Feature Extraction from Texture

The aim of texture feature extraction is to transform the input texture image into a feature vector that describes the properties of a texture, facilitating subsequent tasks such as classification. Among many feature extraction approaches explored in the last two decades especially in the early 1990s and early 2000s, research on texture features focused mainly on four well-established methods of feature description [92]: geometrical methods (e.g. morphological methods), model-based methods (e.g. Markov Random Fields), signal processing methods (e.g. Gabor filters, Laws masks) and statistical methods (e.g. co-occurrence features, Local Binary Pattern). In this section, a number of the most common approaches are described in more detail.

3.4.1 Co-Occurrence Features

The most traditional method for extracting a representative texture feature is the Grey Level Co-occurrence Matrix (GLCM), which is based on second-order statistical texture features. The earliest and most relevant work based on this method was proposed in [111]. GLCM essentially represents the joint probability distribution of a co-occurring pair of pixel values. To extract the GLCM, an image $f(x,y)$ is scanned to count how often pixels whose values are different are separated by a fixed distance d in a certain direction θ , where d is the distance between two pixel positions (x_1,y_1) and (x_2,y_2) and $\theta = 0^\circ, 45^\circ, 90^\circ$ and 135° . From these four directions, four matrixes $P(\theta, d)$ can be derived, which can be expressed as $P(0^\circ, d)$, $P(45^\circ, d)$, $P(90^\circ, d)$, and $P(135^\circ, d)$. Formally, each matrix can be defined as [111]:

$$P(0^\circ, d) = \{P_0(i, j); i \in [0, m], j \in [0, m]\}$$

where each $P_0(i, j)$ is the number of times when

$$f(x_1, y_1) = i, f(x_2, y_2) = j, |x_1 - x_2| = d \text{ and } y_1 = y_2$$

The three remaining matrixes are similarly defined: $P(45^\circ, d)$, $P(90^\circ, d)$, $P(135^\circ, d)$.

Many features can be derived from this matrix; for instance, Haralick [111] proposed fourteen descriptors to represent texture features, which can be categorised into four main groups:

- i. statistical measures including sum average, variance and difference variance
- ii. features extracted from information theory, such as entropy
- iii. visual textural characteristics, including contrast and angular second moment
- iv. correlation-based measures

3.4.2 Local Binary Pattern (LBP) Operator

The Local Binary Pattern (LBP) operator can be seen as a statistical texture descriptor of the characteristics of local structure. It is also regarded as one of the most attractive

operators in image processing and computer vision; as well as efficiently providing the local structures of an input image. LBP is also tolerant of monotonic illumination changes and ease of computation. On the other hand, certain issues related to this operator must be taken into consideration: (a) how to describe different local patterns of textures and how these patterns are extracted; (b) selecting local patterns to form an effective texture descriptor; and (c) representing textures based on the substantial subset of these local patterns. Since 2004, LBP has attracted increasing attention in pattern recognition and computer vision circles, especially in relation to face recognition [112], [84], where LBP descriptor is said to provide complementary information for representing local facial features. For that reason, this subsection examines LBP and its properties in greater detail.

3.4.2.1 A Taxonomy of the LBP Operator

LBP was originally designed by Ojala et al. [88] for the purpose of texture description. It is widely used because of its success as a texture analysis descriptor. As noted earlier, the attractiveness of this approach relates to its robustness against monotonic grey level transformation, and its low computational burden. The basic idea of this approach is to describe a 3×3 neighbourhood around each pixel to generate a binary code; in other words, each pixel is compared to its eight neighbours in a 3×3 window. Any neighbours with values larger than or equal to the central pixel are assigned a value of 1 while neighbours with values lower than the central pixel are assigned a value of 0. In this case, each pixel has a binary number obtained by weighting all binary values around it in an anticlockwise direction. The binary number is then converted to the corresponding decimal value to label the given pixel. Finally, the histogram of labels is utilised as a texture descriptor. Simply put, the 256-bin histogram computed for a region can be used as a texture descriptor. Figure 3.2 illustrates the original LBP operator.

However, the major drawback of the original LBP method is that the small 3×3 neighbourhood could not accommodate dominant features or large-scale structures. It also generates a high dimensionality in data vectors, making it challenging to obtain a reasonable structure. Clearly, if the histograms have too many bins, and the average number of entries per bin is very small, the histograms become sparse and unstable.

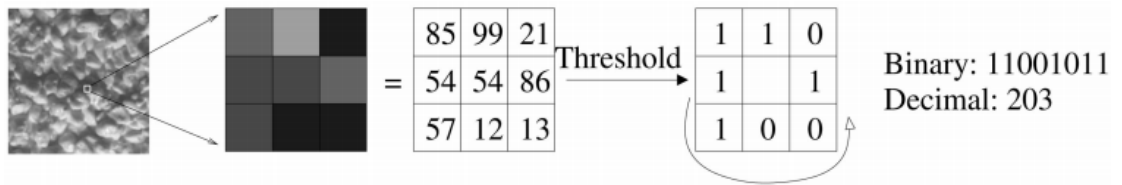


Figure 3.2 Example of the original LBP operator [84]

3.4.2.2 Extension of LBP

In order to deal with texture at different scales, the LBP operator was developed and extended to mitigate the limitations of the original version [113]. Here, local neighbourhoods are defined as a set of points, evenly spaced on a circle centred on the pixel to be labelled, permitting any radius and any number of sampling points. The notation (P, R) is utilised for pixel neighbourhoods, where P represents the number of sampling points and R is the neighbourhood radius. To illustrate, P points around pixel (x, y) lie at coordinates [113].

$$(x_p, y_p) = (x + R \cos(\frac{2\pi p}{P}), y - R \sin(\frac{2\pi p}{P})) \tag{3.1}$$

where $p = 0, 1, 2, \dots, P-1$. If the sampling point does not fall at the centre of a pixel (i.e. at integer coordinates), the pixel value is computed using bilinear interpolation [114]. Figure 3.3 shows three different examples of the extended LBP (ELBP) operator.

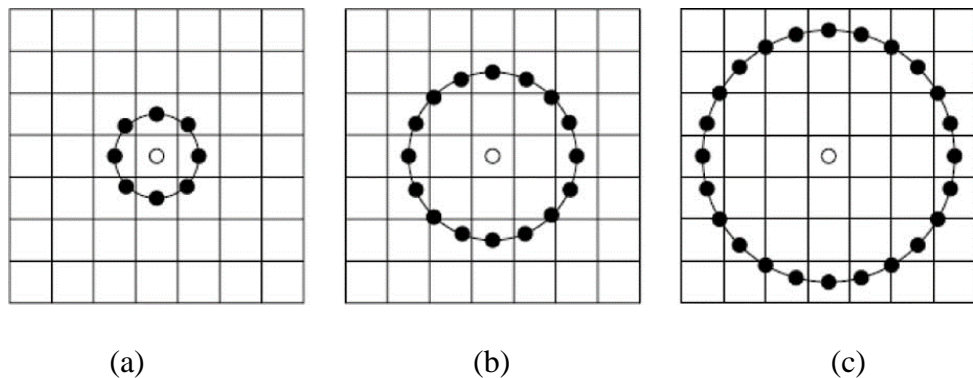


Figure 3.3 Example of the basic LBP operator using different P and R : circular, (a) with $(8,1)$, (b) with $(16,2)$, and (c) with $(24,3)$ neighbourhoods [114]

Formally, given a pixel at (x_c, y_c) , the resulting LBP is formulated as [88]:

$$LBP(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (3.2)$$

where g_c, g_p correspond to the grey scale of the central pixel $c(x_c, y_c)$ and the neighbourhood pixels, respectively. The function $s(x)$ is defined in [88] as follows:

$$s(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases} \quad (3.3)$$

From these definitions, the histogram of LBP labels calculated over a region is exploited as a texture descriptor.

3.4.2.3 Rotation Invariant of LBP Operator

The above descriptions indicate that the LBP can produce 2^p different values, corresponding to the different binary patterns formed using p pixels in the neighbourhood set. However, there is a concern with regard to image rotation; when the image is rotated; the neighbourhood pixels are moved along the perimeter of the circle, resulting in different LBP values. To address this difficulty, a rotation-invariant LBP has been proposed [113], which is formulated as:

$$LBP_{P,R}^{ri} = \min\{ROR(LBP_{P,R}, i)\}, \quad i = 0, 1, \dots, P - 1\} \quad (3.4)$$

where $ROR(x, i)$ performs a circular bitwise right shift, on the p -bit number x , i times [113].

3.4.2.4 Mappings of the LBP Labels: Uniform Patterns

As invariance and robustness to rotation are desirable for many applications, the original LBP was developed to mitigate the effect of local neighbourhood rotation as so-called *uniform patterns* (U). A pattern's uniformity measure is the number of bitwise transitions from 0 to 1 (or vice versa) when the bit pattern is circular. For example,

while the patterns 00000000 (0 transitions), 01110000 (2 transitions) and 11001111 (2 transitions) are uniform, the patterns 11001001 (4 transitions) and 01010011 (6 transitions) are not. The transition or discontinuity should be two transitions at most. In [113], it was observed that, in texture images, uniform patterns account for about 90% of all patterns in an (8, 1) neighbourhood and about 70% in a (16, 2) neighbourhood. The uniform mapping can produce an LBP operator with less than 2^p different labels. For example, the standard LBP with a neighbourhood of 8 pixels generates a dimension of 256, but the uniform LBP (LBP^{U2}) yields only 59 labels, using the same neighbours as with the standard LBP. The $LBP^{U2}_{P,R}$ is defined as [113]:

$$LBP^{U2}_{P,R}(x, y) = \begin{cases} I(LBP_{P,R}(x, y)) & \text{if } U(LBP_{P,R}) \leq 2, I(z) \in 2[0, (P-1)P+2) \\ (P-1)P+2 & \text{otherwise} \end{cases} \quad (3.5)$$

where

$$U(LBP_{P,R}) = |s(g_{p-1} - g_c)| + \sum_{p=1}^p |s(g_p - g_c) - s(g_{p-1} - g_c)| \quad (3.6)$$

If $U(x)$ is less than 3, the current pixel will be labelled by an index function $I(z)$; otherwise, it will be assigned a value of $(P-1)P+2$.

3.4.2.5 Multi-Scale LBP

More structural texture information can be captured by applying a multi-scale LBP. This technique combines the information obtained by N LBP descriptors with varying P and R values. In this sense, each pixel within an image can obtain N different LBP codes. While the LBP codes at different radii are not independent, using multi-resolution analysis often improves the discriminative power of texture features. The multi-scale LBP descriptor was extended for application to face recognition in what is known as the Multi-scale Block Local Binary Pattern (MB-LBP) [115]. The key idea was to compare average pixel values within small blocks; 8 neighbours were considered, generating labels from 0 to 255. For example, using MB-LBP, a block size of 3×3 pixels compares the average grey values at the centre of the block to the

average values of the 8 neighbouring blocks of the same size, resulting in an effective area of 9×9 pixels. The MB-LBP technique is illustrated in Figure 3.4.

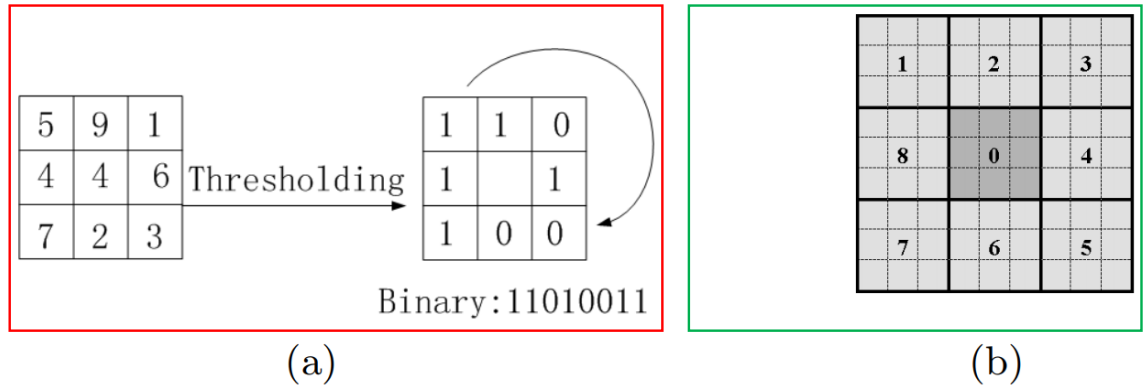


Figure 3.4 Example of MB-LBP showing (a) the original LBP, and (b) the 9×9 MB-LBP descriptor. For each sub-region, the average sum of image intensity is computed and thresholded by the centre block to attain MB-LBP [115]

As indicated above, interest in LBP for texture analysis has been growing since 1996, generating extensive research. Figure 3.5 shows the timeline and milestones for this descriptor over the last two decades.

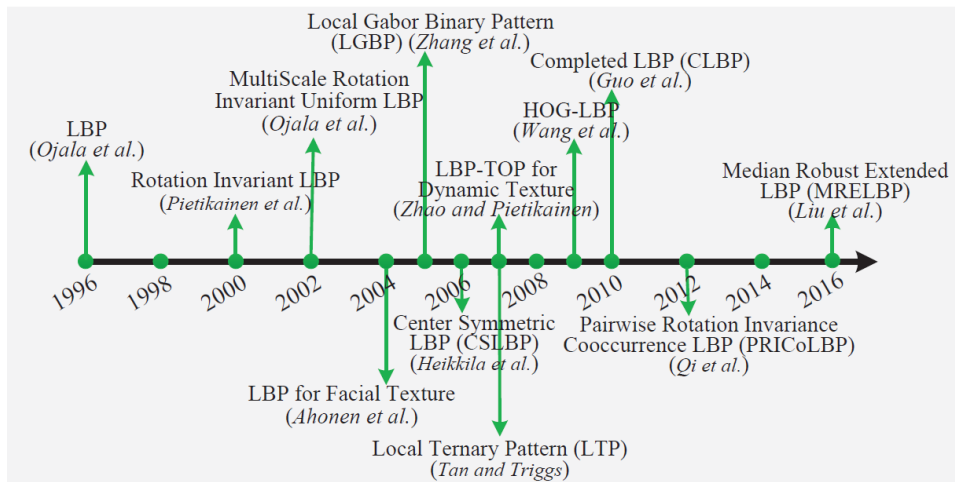


Figure 3.5 Evolution of the LBP descriptor [89]

3.4.3 Gabor Filter-based Texture Features

As a powerful texture analysis tool, Gabor filtering has been widely used for feature extraction, where its function is modified into a linear form. Using this filter, an image is convolved with the filter to generate a response image. It is worth noting that a Gabor

filter is similar to the edge-detection approach. In the signal domain, two methods have been used to describe one-dimensional signals: as a function of time and as a function of frequency. Fourier transforms (or inverse) are used to transform these functions from one method to another; as a result, they produce the same information in different forms. On this definition, the Gabor filter yields a minimal joint-uncertainty of $\Delta t \times \Delta f$ simultaneously in the time (spatial) and frequency domains [116]. It follows that

$$\Delta t \times \Delta f \geq 1/2 \quad (3.7)$$

where Δt is the duration of the signal and Δf is its bandwidth in the frequency domain. In [116], Dennis Gabor proposed that the function can be defined as a complex sinusoidal wave modulated by a Gaussian probability function [116]:

$$\psi(t) = e^{-\alpha^2 (t-t_0)^2} e^{-j2\pi f_0 t + \phi} \quad (3.8)$$

where α is the sharpness (time duration and bandwidth) of the Gaussian probability function, t_0 is the time location of the Gaussian function, f_0 is the frequency of oscillation (or frequency location) and ϕ is the phase shift of the oscillation.

3.4.3.1 1-D Fourier Transform

The one dimensional (1-D) Fourier transform for the Gabor elementary function (Eq. 3.8) can also be defined as

$$\Psi(f) = \sqrt{\frac{\pi}{\alpha^2}} e^{-\left(\frac{\pi}{\alpha}\right)^2 (f-f_0)^2} e^{-j2\pi t_0 (f-f_0) + \phi} \quad (3.9)$$

Eqs. (3.8) and (3.9) show that the Gabor function includes the Gaussian form in both time and frequency domains, which means that the Gaussian can be set at t_0 and f_0 for both time and frequency, respectively. Once the origin-centred filter is located for a convolution task, phase shift and time are removed ($t_0 = 0$, $\phi = 0$). In addition, further

justifications can be provided for defining the Gabor elementary functions using Eq. (3.10):

$$\psi(t) = e^{-\alpha^2 t^2} e^{-j2\pi f_0 t} \quad (3.10)$$

The time duration of the Gabor filter relies on the central frequency f_0 to ensure that, even with different frequencies, Gabor filters are scaled versions of each other, as defined by

$$\alpha = \frac{|f_0|}{\gamma} \quad (3.11)$$

where f_0 is the central frequency of the filter and γ is a parameter that controls the sharpness of the filter (time duration and bandwidth).

Using Eq. (3.9), the maximum response used for a complex signal is given by

$$\sqrt{\frac{\pi}{\alpha^2}} \quad (3.12)$$

In this case, the maximum response should occur at $u = f_0$. Consequently, the 1-D Gabor filter function can be normalised as follows:

$$\psi(t) = \frac{|f_0|}{\gamma\sqrt{\pi}} e^{-\left(\frac{|f_0|}{\gamma}\right)^2 t^2} e^{-j2\pi f_0 t} \quad (3.13)$$

3.4.3.2 2-D Gabor filter

A 2-D Gabor filter in a basic kernel is almost the same as a 1-D Gabor filter; interested readers should refer to the development of the 2-D Gabor filter that began with Granlund in 1978 [117], who introduced a form of general picture operator for solving many properties, such as the octave spacing on the frequency. Works related to this problem are reported by Daugman [118], who developed the uncertainty principle for the two-dimensional plane. Daugman outlined the similarity between the structure based

on the 2D Gabor function and the organisation and characteristics of the mammalian visual system. In the 2-D Gabor filter, the basic parameters are modified as pairs in the x-axis and y-axis. Time t has been replaced, using the spatial coordinates (x, y) in a spatial domain, and the frequency u is replaced by the frequency pair (u, v) in the frequency domain. The 2-D Gabor function can then be defined as [117] [118].

$$\begin{aligned}\psi(t) &= e^{-(\alpha^2 x'^2 + \beta^2 y'^2)} e^{-j2\pi f_0 x'} \\ x' &= x \cos\theta + y \sin\theta \\ y' &= x \sin\theta + y \cos\theta\end{aligned}\tag{3.14}$$

where β is the sharpness of the second Gaussian axis and θ is its orientation. In order to provide a similar shape of filter, sharpness is connected to frequency, which can be achieved by $\alpha = |f_0|/\gamma$ and $\beta = |f_0|/\eta$. The two parameters γ and η control the sharpness of the filter in both the major and minor axes. The 2-D Gabor filter in the spatial domain can be formulated as

$$\begin{aligned}\psi(x, y) &= \frac{f^2}{\pi\gamma\eta} e^{-\left(\frac{f^2}{\gamma^2}x'^2 + \frac{f^2}{\eta^2}y'^2\right)} e^{j2\pi f x'} \\ x' &= x \cos\theta + y \sin\theta \\ y' &= x \sin\theta + y \cos\theta\end{aligned}\tag{3.15}$$

where f is the central frequency of the filter, θ is the orientation angle of the Gaussian major axis and the plane wave, γ is the sharpness along the major axis, and η is the sharpness along the minor axis (perpendicular to the wave). In the given form, the aspect ratio of the Gaussian is $= \eta/\gamma$. As the normalised 2-D Gabor filter function has an analytical form in the frequency domain,

$$\psi(u, v) = e^{-\frac{\pi^2}{f^2}(\gamma^2(u-f)^2 + \eta^2v'^2)}$$

$$u' = u \cos\theta + v \sin\theta \quad (3.16)$$

$$v' = u \sin\theta + v \cos\theta$$

It is clear that the successful use of 2-D Gabor filters depends on optimisation of their parameters. A filter response r for an image function $I(x,y)$ should be calculated at any location (x,y) with the convolution as follows:

$$r_I(x, y; f, \theta) = \psi(x, y; f, \theta) * I(x, y) \quad (3.17)$$

3.5 Feature Classification

The classification task of assigning unknown objects (patterns) to the correct class takes two forms: (1) supervised classification, where the object has a set of data samples with associated labels and class type and (2) unsupervised classification, where the data are not labelled, and the object finds its group by identifying the features that distinguish one group from another. In general, there are two ways of measuring the similarity between feature vectors; the first of these is to measure the distance between the image features, and the second is to measure their similarity. Different distance metrics and other classification approaches have been used for this task. This section reviewed the most commonly used methods of feature classification, including distance metrics and supervised classification.

3.5.1 Distance Metrics

In determining the distance between two patterns, a distance measure is a metric on a feature space used to quantify pattern similarity or dissimilarity; various measures are proposed in the literature [119] [120] [121]. Given two feature vectors a and b , the distance measure d can be called a metric if the following conditions are met:

- Reflectivity: $d(a, b) = 0$, if $a = b$

- Positivity: $d(a, b) > 0$ if a is distinct from b
- Symmetry: $d(a, b) = d(b, a)$
- Triangle inequality: $d(a, c) \leq d(a, b) + d(b, c)$ for every c .

The various distance measures used in different communities include the following:

L1 norm:

$$d(a, b) = \sum_{i=1}^n |a_i - b_i| \quad (3.18)$$

L2 norm:

$$d(a, b) = \sqrt{\sum_{i=1}^n (a_i - b_i)^2} \quad (3.19)$$

Cosine distance:

$$d(a, b) = - \frac{a \cdot b}{\|a\| \cdot \|b\|} = - \frac{\sum_{i=1}^n a_i b_i}{\sqrt{\sum_{i=1}^n a_i^2} \cdot \sqrt{\sum_{i=1}^n b_i^2}} \quad (3.20)$$

Mahalanobis Cosine:

$$d(a, b) = - \frac{m \cdot n}{\|m\| \cdot \|n\|}$$

$$m_i = \frac{a_i}{\sigma_i}, n_i = \frac{b_i}{\sigma_i} \quad (3.21)$$

where a_i, b_i are i th dimension of both vector a and b , and σ_i is the standard deviation of i th dimension in the Mahalanobis space.

An increasing number of machine learning approaches have been used to classify data into groups for biometric purposes, including linear discriminant analysis (LDA) [122] [123] [124], artificial neural networks [125] [126] [127], K-nearest neighbour [128]

[129], support vector machines [130] [131] [132] and sparse representation classifier [133]. Some relevant classifiers used in this research are reviewed in more detail below.

3.5.2 *K*-Nearest Neighbour (*k*-NN)

This classification is the one of the fundamental algorithms in the field of pattern recognition. It is easy and efficient and is also known as the *lazy classification algorithm*. The key principle of *k*-NN is that it compares distance/similarity between template feature samples and probe (or test) samples, where the comparison distance can be any of the distance metrics described in Sub-section 3.5.1. The test samples are then assigned to the nearest labelled samples from the training feature space. Assignment of the test sample is often decided by a *majority vote* rule taking the *k* value into account, where the value of *k* is an integer (usually less than 10). If $k = 1$, the algorithm is called a simple nearest neighbour classifier, and the decision is based on the label of that nearest neighbour's class [128]. If *k* is large, the *k*-NN yields smooth boundaries [129].

Cover and Hart [134] proposed that the error rate of the 1-nearest-neighbour classifier is never more than twice the Bayes rate for the same data. Consequently, the main attraction of this algorithm, along with its asymptotic performance, is that it optimises selection of only one parameter *k*. A full account of the theoretical basis of the *k*-NN classifier has been introduced in [96] and [98]. Despite the popularity of the *k*-NN classifier, it has some limitations; for example, it is sensitive to local distribution of the training sample in the feature space, which may lead to unstable performance. The density of the data clusters also affects performance and may lead to wrong decision making; some *k*-NN algorithms have therefore used distance *d* as a weight ($1/d$) to improve decision making [128].

3.5.3 Support Vector Machines (SVM)

Support Vector Machines (SVMs) are the most powerful classification methods and have been employed in a wide range of classification and regression problems, including face detection [130], face verification [131] and gender classification [132]. SVM is a supervised learning method that uses so-called 'support vectors' to build a

model for classification or regression. The basic idea of SVM was proposed in [135], and the algorithm was developed in [136]. The SVM classifier integrates the notation of supporting vectors and kernel tricks. Given a set of training samples, it aims to determine a maximal separating hyperplane between two classes so that classification error is minimized. As the available data are not always linearly separable in the feature space, the data feature space can be mapped into a space with higher dimensionality to simplify separation in that space [137]. Mapping of the feature space onto higher spaces is the so-called *kernel trick*. SVM is described in more detail in [138].

Given training samples (x_i, y_i) , where $x_i \in R^n$, $y_i \in \{-1, +1\}$ for $1 \leq i \leq l$, and l is the number of samples, the hyperplanes can be expressed as

$$H_1: w^T x_i - b = +1 \quad (3.22)$$

$$H_2: w^T x_i - b = -1$$

where b is the offset of the hyperplane from the origin and w is the n dimensional vector perpendicular to the separating hyperplane (see Figure 3.6). Based on Eq. (3.22), the solution should meet the following constraints:

$$\begin{cases} w^T x_i + b \geq +1 \quad \forall y_i = +1 \\ w^T x_i + b \geq -1 \quad \forall y_i = -1 \end{cases} \quad (3.23)$$

These two constraints are combined into one set of inequalities:

$$y_i(w^T x_i - b) \geq 1, i = 1, 2, \dots, l \quad (3.24)$$

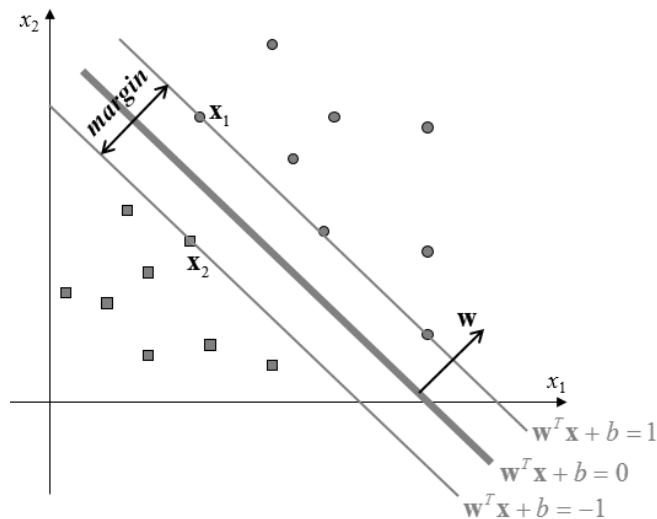


Figure 3.6 Separation of the data into two classes

If the data are linearly separable, the separation of their hyperplane is formulated as:

$$\omega^T x_i - b = 0$$

where the distance between the canonical hyperplanes and the separating hyperplane is calculated as:

$$\frac{1}{\|\omega\|}$$

In this case, optimising the separating margin is equivalent to maximising the distance between H_1 and H_2 . The largest distance between H_1 and H_2 is calculated as:

$$\frac{2}{\|\omega\|}$$

The SVM learning problem (i.e. optimisation problem) is expressed as follows:

$$\text{minimise } \frac{1}{2} \|\omega\| \quad \text{subject to } \omega^T x_i - b \geq 1 \quad (3.25)$$

Saddle points of Lagrange's function can tackle the optimisation problem as follows:

$$\text{minimise } L(\omega, b, a) \equiv \frac{1}{2} \|\omega\|^2 - \sum_{i=1}^l \alpha_i [\omega^T x_i - b] - 1 \quad (3.26)$$

where $\{\alpha_i : 1, \dots, l; \alpha_i \geq 0\}$ are Lagrangian multipliers; ω, b , and $\alpha_i \geq 0$ can minimise L and the dual form of the Lagrangian, which can be formulated as follows:

$$\begin{aligned} \text{maximise } L_D &= \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j x_i^T x_j & (3.27) \\ \text{subject to } & \sum_{i=1}^l \alpha_i y_i = 0, \alpha_i \geq 0 \forall_i \end{aligned}$$

If the data are not linearly separable, the technique should be changed. The positive slack variables $\{\zeta_i, i = 1, \dots, l, \zeta_i \geq 0\}$ are introduced into the original constraints, with additional penalty value C for points that cross the boundaries to consider the classification error.

$$\begin{aligned} \text{minimise } & \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^l \zeta_i & (3.28) \\ \text{subject to } & y_i(\omega^T x_i - b) \geq 1 - \zeta_i, \zeta_i \geq 0 \end{aligned}$$

To exclude ω, b , and ζ , the linear technique can be applied; the dual form of Lagrange's function is expressed as:

$$\begin{aligned} \text{maximise } L_D &= \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j x_i^T x_j & (3.29) \\ \text{subject to } & \sum_{i=1}^l \alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq C \forall_i \end{aligned}$$

Nonlinear classifiers were proposed in [139] by using the kernel trick. The kernel function is defined as a map that transforms data from the original space whose data are not linearly separable to a higher dimensional space to make data linearly separable. The kernel function can be formulated as $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$. The discriminant function is $g(x_i) = \omega^T \phi(x_i) + b$, and the corresponding dual form replaces $x_i x_j$ in [140] by $K(x_i, x_j)$:

$$\begin{aligned} \text{maximise } L_D &= \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \phi(x_i)^T \phi(x_j) & (3.30) \\ \text{subject to } & \sum_{i=1}^l \alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq C \forall_i \end{aligned}$$

The most common kernel functions used to reliably build SVM classifiers are:

Linear:

$$K(x_i, x_j) = x_i^T x_j \quad (3.31)$$

Polynomial:

$$K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0 \quad (3.32)$$

Radial basis function (RBF):

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (3.33)$$

Sigmoid:

$$K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r) \quad (3.34)$$

where γ , r and d are kernel parameters.

3.5.4 Sparse Representation Classifier (SRC)

The sparse representation classifier (SRC) was introduced in [133], achieving high performance for face recognition. Given a set of labelled training samples from K

classes, the task is to determine the class to which a new unseen test sample belongs. Let n_i be the training samples from the i -th class; all samples are arranged as columns.

$$A_i = [v_{i,1}, v_{i,2}, \dots, v_{i,n_i}] \in R^{m \times n} \quad (3.35)$$

Each column $v_{i,j}$ in matrix A_i is the vectorised intensity image. In this case, given enough samples from the i -th class, any new test sample \mathcal{Y} from the same class is approximated to lie in the linear span of the columns of matrix A_i . This can be formulated as follows:

$$\mathcal{Y} = \sum_{j=1}^{n_i} \alpha_{i,j} v_{i,j} \quad (3.36)$$

where $\alpha_{i,j} \in R, j = 1, 2, \dots, n_i$. The matrix A is built by concatenating the training samples from all classes:

$$\begin{aligned} A_{m \times n} &= [A_1, A_2, \dots, A_k] \\ &= [v_{1,1}, v_{1,2}, \dots, v_{1,n_1} \mid \dots \mid v_{k,1}, \dots, v_{k,n_k}] \end{aligned} \quad (3.37)$$

where $n = \sum_{i=1}^k n_i$ is formulated as:

$$\mathcal{Y} = Ax \in R^{m \times n} \quad (3.38)$$

$$x = [0, \dots, 0, \dots, \alpha_{i,1}, \dots, \alpha_{i,n_i}, \dots, 0, \dots, 0]^T \quad (3.39)$$

x in (Eq. 3.39) is the a coefficient vector that consists of zero except those corresponding to the i -th class. The objective of a given new test sample \mathcal{Y} and matrix A is to obtain an x that is informative in terms of recognition of test samples. The sparse representation approach looks for the sparsest solution to the linear system ($\mathcal{Y} = Ax$); sparser is better in identifying the unknown test sample, solving the following optimization problem:

$$(l^0) : x_0^\wedge = \arg \min \|x\|_0 \quad \text{subject to } Ax = \mathcal{Y} \quad (3.40)$$

where $\|x\|_0$ represents the l^0 -norm, which counts the number of non-zero entries in a vector. Where the solution of x_0 is sparse enough, the solution l^0 -minimisation problem equates to the following l^1 -minimisation problem:

$$(l^1) : x_1^{\wedge} = \arg \min \|x\|_1 \quad \text{subject to} \quad Ax = \mathcal{Y} \quad (3.41)$$

Using (3.41), the residual r between the training sample and the test sample can be formulated as [133]:

$$r_i = \left\| \mathcal{Y} - \sum_{j=1}^k x_{1,i,j} u_{i,j} \right\|_2 \quad (3.42)$$

The test sample is assigned to the subject with the smallest residual.

3.6 Biometric Information Fusion Strategies

Biometric information from different sources may lead to error in recognition systems, mainly because of the limited ability to represent subjects, as well as practical problems such as noisy sensor data or indistinct information. Integration of different sources of biometric information may be the best option to improve the accuracy, robustness and efficiency of biometric systems and to also reduce recognition error rates. Fusion of biometric data can be implemented at various levels:

- i. Sensor level: raw data acquired from multiple sensors are usually integrated for new data feature extraction.
- ii. Feature level: feature sets extracted from multiple sources can be fused to generate new features to represent the subject.
- iii. Match score level: output scores of multiple classifiers are fused to form a single score on which to make the final decision.
- iv. Rank level: this level is often related to identification systems where each classifier generates a set of scores for enrolled subjects (higher rank indicates best match).

- v. Decision level: using matching scores, the claimed identity is determined as genuine or impostor (i.e. verification system accept or reject).

For interested readers, all of these techniques are explained in more detail in the literature [141] [142] [143] [144]. For more interpretation, fusion-based biometric systems can be implemented at the following levels.

3.6.1 Sensor-level Fusion

As mentioned above, raw data acquired from multiple sensors are combined at sensor level; for example, face images captured from multiple cameras can be combined to form a single face image. The data obtained from these sensors must be compatible in order to be fused at this level.

3.6.2 Feature-level Fusion

In feature-level fusion, feature sets originating from multiple biometric systems or from multiple algorithms are fused to form a single feature vector. Two different techniques can be used to fuse feature vectors: if the feature vectors are homogeneous, a single feature vector can be formed using a weighted average of subject feature vectors; if the feature vectors are non-homogeneous or from different modalities, a single feature vector can be formed by concatenation.

3.6.3 Score-level Fusion

Each matcher provides a matching score showing the proximity of the reference feature vector to the probe. The scores are fused to generate a new matching score that can be used to render an identity decision. Because of its performance and simplicity, this strategy is the most commonly discussed level of fusion. Certain common rules are usually applied to generate a final score, including *Maximum*, *Minimum*, *Sum*, *Product*, and *Mean* [145]. In some cases, score normalisation techniques are applied to transform scores into a common domain. The various score normalisation techniques reported in the literature [145] [146] include *Min-Max*, *Z-Score*, *Tanh*, *Median-MAD* and *Double-sigmoid*.

3.6.4 Decision-level Fusion

In this approach, decisions produced from multiple classifiers are combined using different techniques including “AND” and “OR” rules [147], *majority voting* [148], *weighted majority voting* [149] and *Bayesian decision fusion* [150].

3.7 Measuring Biometric System Performances

In the academic literature, the performance of biometric systems is typically quantified in terms of two measures: cumulative match characteristic (CMC) and receiver operating characteristic (ROC). Both of these measures are plotted graphically as curves. The CMC curve commonly represents the degree of change in recognition accuracy for the top k ranked similarity scores. More generally, it is used to evaluate performance in identification mode. The CMC plot shows the probability of correct identification (y-axis) against the number of individuals (x-axis). Although rank-one recognition accuracy is often preferred, some studies consider the first ten or twenty similarity scores returned [151].

ROC curves are commonly used to evaluate performance in verification mode. A biometric verification system is subject to two types of error: False Acceptance (FA) and False Rejection (FR). FA occurs when two samples obtained from two different individuals are assigned to one individual; FR occurs when two samples obtained from the same individual are assigned to two different individuals. False Acceptance Rate (FAR) and False Rejection Rate (FRR) are calculated as follows:

$$FAR = \frac{\text{Number of false acceptances}}{\text{Total number of impostors}} \quad (3.43)$$

$$FRR = \frac{\text{Number of false rejections}}{\text{Total number of genuines}} \quad (3.44)$$

ROC curves can provide statistical measures such as equal error rate (EER), area under the curve (AUC), and genuine acceptance rate (GAR). These curves express the trade-off between GAR (1-FRR) (plotted on the y-axis) and FAR (plotted on the x-axis), using a set of threshold values (θ). For a given a value of θ , the EER is the rate where

$FAR = FRR$. The sum of both errors (FAR and FRR) represents the system's total error rate (TER). The verification rate ($1-FRR$) at $FAR = 0.1\%$ is commonly used to represent system accuracy.

3.8 Summary

This chapter covered the most significant areas of immediate relevance. After introducing the challenges of texture analysis techniques and related issues, well-known methods for extracting texture features were discussed, including Local Binary Pattern (LBP) and Gabor wavelets filters. Texture feature classification approaches were also outlined, followed by different schemes for fusing biometric information. Finally, the different ways for evaluating the performance of biometric recognition systems were presented. The next chapter provides a detailed description and discussion of the proposed new technique for skin pixel detection and its purity analysis.

Chapter 4

Detecting Facial Skin Regions for Biometric Use

4.1 Introduction

This chapter describes the design, implementation and evaluation of a technique for skin detection to be incorporated into biometric systems. This incorporation reduces the challenges associated with human interactions and provides the potential for automation of skin-based biometric recognition. Following an overview of existing approaches to skin detection, the chapter outlines various skin colour models and their limitations. The chapter also discusses the databases used for the evaluation, the generation of ground truth by subjective observations, as well as skin region purity analysis. Finally, the chapter reports the qualitative and quantitative results obtained using the proposed framework.

The chapter is organised as follows. Section 4.2 describes the challenges for skin detection, followed by a brief but precise review of the relevant literature in Section 4.3. Sections 4.4 and 4.5 describe the general framework for the proposed skin detection technique and skin quality assessment. The experimental setup and results are presented in Section 4.6 and Section 4.7. The chapter concludes with a summary in Section 4.8.

4.2 Challenges for Skin Detection

Skin colour detection is utilised in a wide range of applications and involves both image processing and computer vision. Existing techniques for face detection [152], gesture recognition [153] [154] and tracking require a well-designed algorithm for skin detection [155] [156]. Although the skin detection problem is well known to researchers, it is difficult to find a single algorithm for skin detection that will provide accurate results under different conditions [157]. At the main time, the simplicity of skin colour detection technique has attracted researchers for more investigation.

Most of the existing research on skin detection focuses on skin information and segmentation of skin colour pixels [158], [159]. Skin colour in an image is dependent to a number of factors. For example, camera characteristics: the same person's skin colour may look different when captured by different cameras, as camera sensor characteristics determine the captured image's properties. Illumination: distribution of light changes with ambient environmental variables (indoor, outdoor, shadows etc.), thus, creating the so-called 'colour constancy problem' [160]. Individual characteristics: an individual's skin colour may be affected by factors such as age, gender or body part exposure. Ethnicity: skin colour varies across different regions, as in Asian, African and Caucasian variants. Other factors: artefacts such as makeup or glasses impact on skin colour, as do hairstyle, background and shadows. All these factors fall out of the scope of the present thesis. Thus, interested readers are referred to [158] [159] and [160].

Any skin detection technique should endeavour to minimise the impact of these parameters for satisfactory result.

4.3 Previous Works on Skin Detection

There are two types of skin detection techniques: pixel-based and region-based technique. Pixel-based technique is to classify each pixel individually whether skin or non-skin from its neighbour. The region-based detection technique requires additional characteristics such as texture and intensity. While Skin detection based on colour belongs to pixel-based category, the task here associated with skin detection is to assign pixels to one of two categories: skin pixel and non-skin pixel. By exploring skin

detection techniques for use in this study, several colour spaces were mainly reviewed. Studies of skin detection based on colour commonly use one of three main methods [161]:

- parametric skin distribution modelling, including Gaussian and Gaussian mixture models;
- non-parametric skin distribution modelling, including lookup tables and Bayes skin probability maps;
- cluster colour space, where skin regions are treated as cluster pixels

While skin detection remains an integral part of face recognition systems, the appearance of the facial skin can change dramatically for various reasons, including illumination, pose, expression and image quality. Skin detection involves the capture of image pixels and regions that constitute skin tone colour. In studies of skin colour pixel detection [159], [162], different colour spaces are used for this task. For example, the early work introduced by Sobottka et al. [163] developed skin detection method based on HSV colour space. Hsu et al. [164] adopted various thresholds by partitioning HIS colour space into three zones to classify skin pixels. Zaqout et al. [165] used the RGB colour space to segment skin regions, and Yang et al. [166] proposed an approach based on the YCbCr colour space to detect skin pixels, using neural network classification. Abdul Rahman et al. [167] employed a skin colour model based on three colour spaces—RGB, HSV (using only H values) and YCbCr—to detect human faces. Sawicki et al. [168] proposed a thresholding method for detecting skin colour, using CMYK colour space.

For present purposes, the cluster colour space method seems preferable because of the algorithm's simplicity and low computational complexity. Human skin colour is distinct from other objects and can be recognised for the purposes of classification. It is clear that explicit thresholding facilitates construction of fast methods. A commonly used method is to define a skin colour decision boundary for different colour space components, specifying a range of threshold values for each colour space component. If the image pixel value falls within the predefined range(s), it is considered a skin pixel. The proposed framework involves methods of this kind.

The present framework utilises different colour spaces to differentiate between skin region pixels and other objects within the facial image for biometric purposes. As use of a single colour space to detect skin regions may fail to overcome the issues referred to above, deploying more than one colour space can help to eliminate some of these issues and so improve detection performance. In the proposed framework, three colour spaces—RGB, HSV, and YCbCr—are applied in different ways to assess the purity of the skin region of interest (ROI). Although the RGB colour space is widely used in image segmentation methods, it does not separate luminance and chrominance information; HSV and YCbCr enable separation of the requisite colour information for skin pixel detection.

4.4 A General Framework for Skin Purity Analysis

The motivation for developing the proposed framework is that it can be used as a prior biometric process to automate skin-based biometric systems. The general structure of the framework is set out in Figure 4.1, showing the stages of skin detection prior to use in a skin-based biometric system.

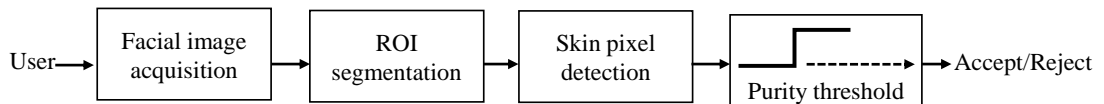


Figure 4.1 General structure of the proposed skin detection framework

The technique proposed here focuses on the assessment of facial skin images for biometric purposes. Four facial skin regions were chosen for evaluation: forehead, right cheek, left cheek and chin. The objective of the proposed system is to automate the process of deciding whether the target region contains sufficient skin. As the biometric system developed in this research is based on skin regions from partially occluded facial images, other global facial features (such as eyebrows, eyes, nose and mouth) would not be captured, and this is one reason for partitioning the face image into four skin regions. It is also likely that skin content in one of these regions can be captured even when the whole face is not visible, enabling the system to decide whether that skin region can be

used for biometric purposes. For example, in video surveillance frames, the captured image should contain at least one of these facial skin regions, and it is worth assessing each region independently for biometric information.

4.5 Skin Colour Schemes

The extracted ROI is automatically assessed to determine whether it can be used for biometric purposes. Clearly, any ROI that is unduly corrupted by artefacts—whether natural (e.g. hair, moles) or artificial (e.g. headgear, glasses)—will be considered invalid for biometric analysis. Unlike other studies based on the use of features encoded in skin (moles, birthmarks, etc.) as discriminatory biometric features, or where skin regions are manually chosen, the present goal is to automate assessment of pure skin as a biometric source, as moles and other artefacts are not universal across all individuals. Automation not only enhances efficiency but also ensures the accuracy and consistency of large-scale processing.

In this implementation, colour modelling was utilised to assess the purity of the extracted ROI. While skin tone typically occupies a compact region within the colour space, skin colour can change for the reasons noted above (e.g. variations in illumination, camera sensors, gender, age, ethnicity). As demonstrated earlier, among the many reasons for using colour spaces to distinguish skin pixels from other elements in a facial image, skin colours are computationally robust against challenges like partial occlusion, scaling, pose and rotation.

As no single colour space is likely to overcome the issues referred to in Section 4.2, applying more than one colour space helps to mitigate such issues and improves detection performance.

Additionally, morphological operations (*‘opening’* followed by a *‘majority’* operator and a 3x3 square structuring element) were applied to refine skin regions and to eliminate incorrectly marked small groups of pixels (due to noise etc.) within the ROI. Figure 4.2 specifies the algorithmic flowchart for the proposed skin purity analysis technique.

For this study, two rules derived from the three colour spaces are used to identify skin pixels in the extracted ROI. These are subsequently combined to assess whether the ROI contains adequate skin to merit subsequent processing. While the first model uses the RGB colour space, the second uses CbCr and H values from the HSV and YCbCr colour spaces. These rules are described as follows:

Scheme 1: This scheme utilises only the RGB colour space as originally proposed in [169]. In this model, a pixel is marked as pure skin if its RGB values satisfy the conditions in (4.1) below:

$$\text{Pixel} = \begin{cases} \text{skin,} & \text{if } R > 95 \text{ and } G < 40 \text{ and } B > 40 \text{ and} \\ & (\max(R,G,B) - \min(R,G,B)) > 15 \text{ and} \\ & |R - G| > 15 \text{ and } R > G \text{ and } R > B \\ \text{non-skin,} & \text{otherwise} \end{cases} \quad (4.1)$$

Scheme 2: This scheme detects skin pixel values within the region of interest by combining two colour spaces: HSV and YCbCr. It is clear that any other colour space can be obtained from RGB. Generally speaking, the transformation between RGB and YCbCr is basically linear while the transformation between RGB and HSV is nonlinear. To identify skin pixels using this scheme, the RGB values for each pixel are first converted to HSV and YCbCr components. Once these values have been computed for each pixel, they can be classified as skin or non-skin by comparing the Cr, Cb and H values against a set of threshold values.

Scheme 2 proposed here is basically based on pixel colour analysis. Experimentally, it is found that as skin pixel values are expected to fall within the ranges of $140 \leq Cr \leq 170$, $140 \leq Cb \leq 198$ and $0.01 \leq H \leq 0.10$, values for H (hue), Cb and Cr are thresholded. Any pixel values falling outside these ranges are classified as non-skin pixels. The adaptive skin colour scheme uses trial and error approach to choose threshold values for lower and upper bounds of the three components (Cr, Cb, and H). For this scheme, the thresholds have finally been formulated as:

$$\text{Pixel} = \begin{cases} \text{skin,} & \text{if } 140 \leq Cr \leq 170 \text{ and} \\ & 140 \leq Cb \leq 198 \text{ and} \\ & 0.01 \leq H \leq 0.10 \\ \text{non-skin,} & \text{otherwise} \end{cases} \quad (4.2)$$

Both model schemes yield masks that identify which pixels within the ROI are to be classified as skin. If the number of skin pixels within the mask is greater than a predefined threshold of skin purity assessment, the skin ROI is considered pure.

Scheme 3: This scheme combines components from Scheme 1 and Scheme 2. If either Scheme 1 or Scheme 2 classifies the ROI as a pure skin region, the ROI is accepted for biometric analysis.

4.6 Experimental Setup

This section provides more detail about the databases and the approach used to determine the ground truth for assessment of the skin detection technique.

4.6.1 Databases

Two databases were used to evaluate the proposed framework. The Skin Segmentation dataset [170] was used to evaluate the skin pixel detection technique, and XM2VTS [171] was used for skin regional purity assessment.

4.6.1.1 The Skin Segmentation Dataset

The first set of experiments (related to skin/non-skin detection) were conducted using the Skin Segmentation dataset [170]. This skin/non-skin dataset contains 245,057 pixel samples. 50,859 are skin pixel samples and 194,198 are non-skin pixel samples in the RGB colour space. The data were generated from two facial image databases called PAL and FERET. The dataset contains skin pixel samples from facial images representing various ages (young, middle, and old) and racial groups (White, Black, and Asian), including both males and females.

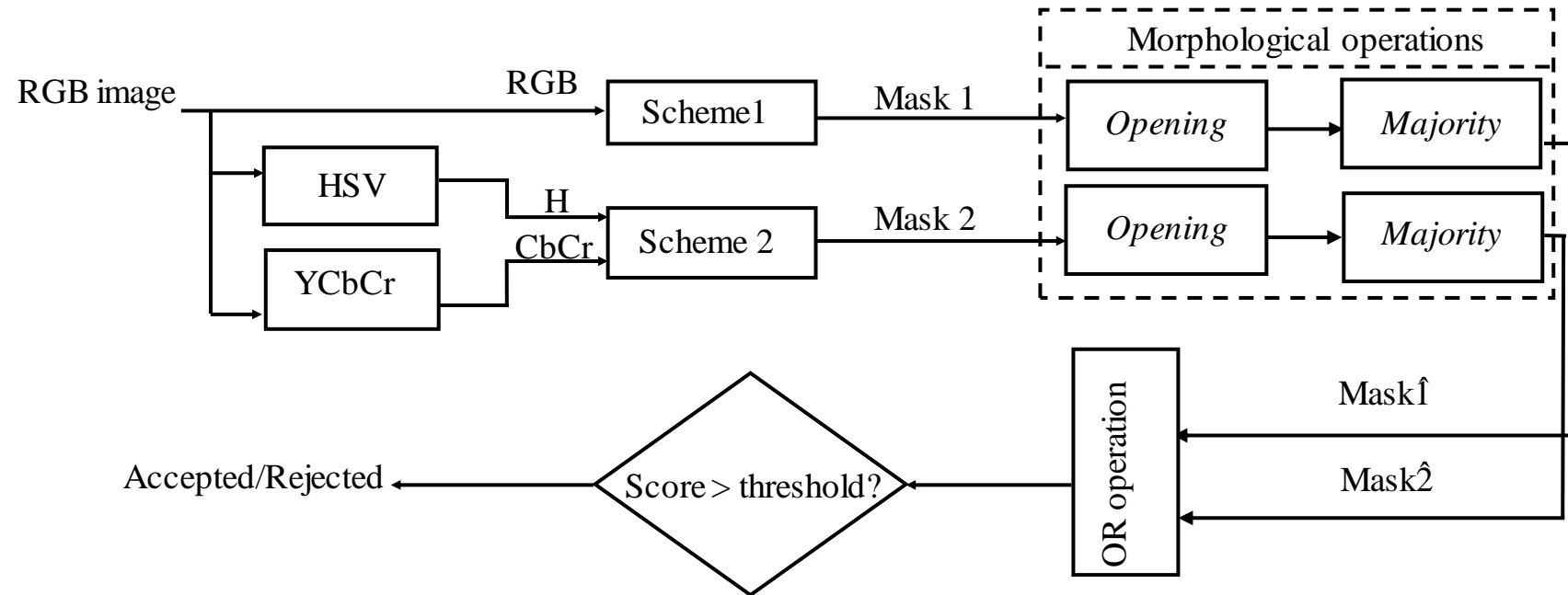


Figure 4.2 Proposal for a general skin detection system. The RGB skin image input is transformed to HSV and YCbCr. While skin modelling Scheme1 uses only one colour space (RGB), skin modelling Scheme2 combines two further colour spaces (HSV and YCbCr). Skin-modelling Scheme2 is based on H, Cb and Cr values

4.6.1.2 XM2VTS Database

The validity of the proposed framework was tested using the XM2VTS multimodal database [171], comprising colour stills, video sequences and speech recordings taken from 295 individuals of both genders and of various ethnic origins. Four frontal face images of each subject were captured during four sessions over a period of five months, in a controlled environment with uniform illumination and a blue background. The database includes 1,180 images (295 individuals x 4 images), each measuring 720×567 pixels.

4.6.2 Generating the Ground Truth for Purity Analysis

Judgments based on subjective observations are widely accepted and used in the literature on image processing and machine learning [172]. It is equally clear that the evaluation of any skin detection algorithm depends on the availability of skin image samples that are reliably and appropriately tagged with an accurate ground truth label. Achieving a consistently high level of precision in labelling skin images is likely to prove challenging, especially for skin/non-skin boundaries. This problem is not unusual in machine learning and image processing, as for instance when attempting to define and quantify concepts of sadness or happiness when processing facial expressions. Making assumptions based on self-assessments may offer one means of generating an acceptable ground truth, as experienced viewers can provide useful measures, especially when evaluating a number of images. On that basis, all images in the XM2VTS database [171] were manually labelled with multiple observers' subjective judgements about the ROI purity. Although most observers had experience in working technically with image processing, it was difficult to define the concept of skin region purity, and the observers were asked to choose the best definition. For the purposes of objective assessment, 79 images were selected. The 79-selected image included 19 images from forehead and 20 images from other regions (right cheek, left cheek and chin region), with different noise sources (hair, glasses, reflections etc.). The overall decision integrated the judgments of 15 evaluators, using majority rule for final definition of skin purity. The 15 researchers worked in relevant fields, including machine learning, image processing and biometric technology. Finally, the stability of the evaluators' annotations




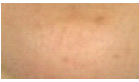
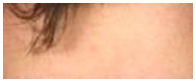


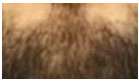
and their decisions were also studied and analysed. Table 4.1 summarises the outcome of this manual ground truth generation for the four facial skin regions in the XM2VTS database images. It should be noted here that the Chehra software [20] (this approach is explained in Section 5.2 with more details) failed to locate landmarks in 52 of the 1,180 XM2VTS images. Of the remaining 1,128 images, all four images were marked for only 274 individuals (out of 295). These 1,096 images (274 x 4) were analysed for the purposes of this study.

Table 4.1 Ground truth of skin and non-skin images of the four facial regions from XM2VTS database

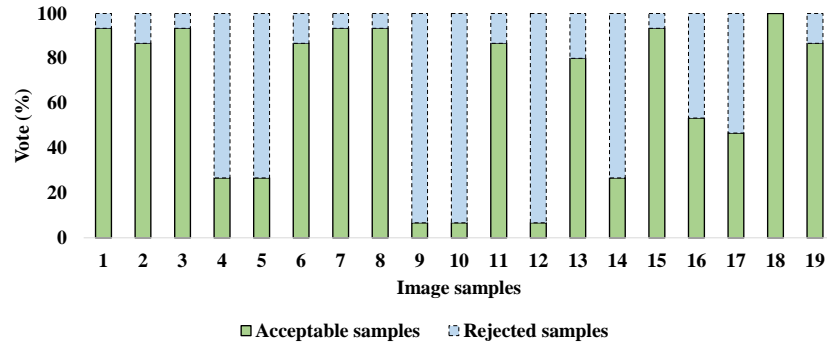
Decision	No. of images per region in ground truth (manual)			
	Forehead	Left Cheek	Right Cheek	Chin
Acceptable	700	1025	1039	997
Non-acceptable	396	71	57	99

As shown in Table 4.1, of the 1,096 extracted ROIs for the forehead, 700 were classified by the human operators as pure skin regions while ROIs for the left and right cheek numbered 1025 and 1039, respectively. For the chin, 997 ROIs were classified as pure skin. The remaining images were considered corrupted by various artefacts. Table 4.2 shows some sample ROIs with assigned categories.

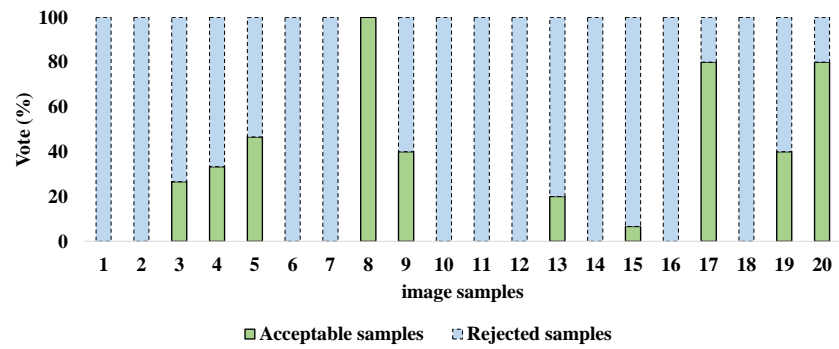
Table 4.2 Some samples of ROIs graded as accepted and rejected by the evaluators (ground truth)

Decision	ROI			
	Forehead	Left Cheek	Right Cheek	Chin
Acceptable				
Rejected				

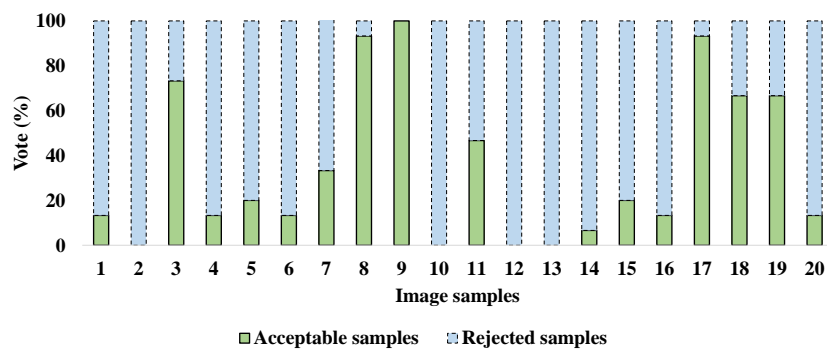
A case study of objective observations judging the suitability of sample skin images for skin purity ROI is analysed. Figure 4.3 shows the decisions made by the observers.



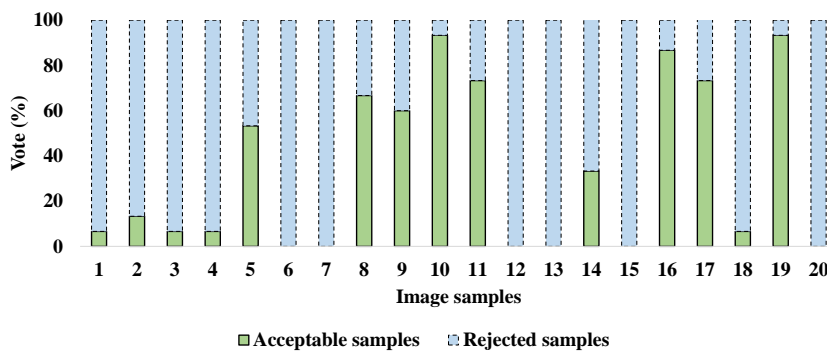
(a) Forehead region



(b) Right cheek region



(c) Left cheek region



(d) Chin region

Figure 4.3 Manual labelling of the four facial skin regions by 15 observers to create a ground truth from the XM2VTS database

Although it is very hard to judge whether the tested image sample contains sufficient skin information for biometric analysis, an overwhelming majority of the observers agreed in their final decision. The annotations of the different evaluators and their discriminative power are clear evidence of the stability of their decisions made.

4.6.3 Automated Skin Purity Analysis

Purity of the ROI may be degraded by the presence of hair or artefacts such as glasses, makeup, and reflections caused by lighting. These factors affect the accuracy of skin biometric performance. Clearly, a “usable” ROI should have a large majority of pure skin pixels at the cost of some noise in the image (e.g. hair, artefacts), using a predefined threshold. Here, based on the idea described in [173], the threshold for measuring goodness of ROI was set in the 95–99% range, allowing a small amount of non-pure skin pixels in the ROI (termed false acceptance). It was determined empirically that skin detection performance is not affected by items in this range.

4.7 Experimental Results and Discussion

The main objectives of the present study were to automate the skin-based biometric system and to extend the population coverage of biometric processing as far as possible. This section discusses the outcome of the proposed framework and compares it to the ground truth generated by human volunteers. Skin images categorised as satisfactory are recommended for further biometric use in scenarios such as identification or verification.

4.7.1 Performance of the Proposed Skin Pixel Detection

This sub-section only studies the performance of the proposed skin pixel detection scheme (Scheme 2). The skin pixel detection technique was proposed to automatically recognise skin and non-skin pixels in each region of interest. The scheme was developed on the basis of pixel colour analysis. The RGB colour space is converted to HSV and YCbCr colour spaces. Pixel values are thresholded where skin pixel values are expected to fall within the quantified ranges (refer to Eq. 4.2). All pixel values falling

outside these ranges are considered non-skin. The approach performance was measured in terms of True Positive Rate (TPR) and True Negative Rate (TNR), which are defined as follows.

$$\text{TPR} = \frac{\text{No. of skin pixels correctly classified}}{\text{Total no. of skin pixels}} \quad (4.3)$$

$$\text{TNR} = \frac{\text{No. of non skin pixels correctly classified}}{\text{Total no. of non skin pixels}} \quad (4.4)$$

Using the above parameters, experiments were performed using the Skin Segmentation dataset [170]. Table 4.3 summarises the results, including the efficiency of the proposed technique and comparative results for other existing skin pixel detection techniques that used the same dataset. Non-skin detection was more accurate (99.91%) than skin detection (93.97%), indicating that the proposed technique is robust for detection of non-skin pixel values. However, some skin and non-skin pixels are clearly being misclassified; almost 6% of skin pixel values were falsely rejected, and more investigation is needed. In the proposed system, any ROI containing skin pixels below a predefined threshold would be rejected to eliminate distortions caused by artefacts such as spectacles, reflections, or shadows.

Table 4.3 Performance of skin/non-skin detection technique compared with other published results

Method	Technique	TPR	TNR
Bhatt et al. [170]	Fuzzy Decision Trees	98.09	92.51
Çatak [174]	MapReduce-based distributed Ada Boosting of extreme learning machine (ELM)	99.56	N/A
<i>Proposed technique (Scheme 2)</i>	Thresholding Cb, Cr & hue (H) values $140 \leq Cr \leq 170$ and $140 \leq Cb \leq 198$ and $0.01 \leq H \leq 0.10$	93.97	99.91

Note that the threshold of skin regional purity assessment for this scheme and its impact on skin-based biometric system are investigated separately in Section 5.6.

4.7.2 Performance of Skin Detection Regional Purity Analysis

This experiment was conducted on the XM2VTS database to assess all colour schemes with their performances on skin region of interest. As demonstrated in Section 4.5, three skin pixel detection techniques were investigated, using different colour spaces. The three schemes were termed Scheme 1 (using only the RGB colour space), Scheme 2 (using Cb, Cr and H values from two different colour spaces) and Scheme 3 (combining Schemes 1 and 2). These schemes were therefore used to assess and validate the purity of skin regions extracted from facial images using XM2VTS database. Where the proportion of skin pixels was higher than a chosen threshold, the ROI was deemed to be of adequate purity. The measures used to evaluate the performance of the proposed framework were set out in metrics numbered from 4.5 to 4.9. Different terms are formulated which can be defined as follows: TP refers to cases where colour models correctly classify an acceptable purity skin ROI, and TN refers to cases where a corrupted region is correctly classified. FP refers to cases where the corrupted region is classified as adequate, and FN refers to cases where an adequate region is classified as corrupted. All measurements used here are formulated as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (4.5)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4.6)$$

$$\text{False Positive Rate (FPR)} = \frac{\text{FP}}{\text{FP} + \text{TN}} \quad (4.7)$$

$$\text{False Negative Rate (FNR)} = \frac{\text{FN}}{\text{TN} + \text{FP}} \quad (4.8)$$

$$\text{F - Measure} = 2 \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4.9)$$

As noted above, optimisation for the best threshold (θ) can be obtained using different thresholds. Figure 4.4 shows ROC curves for five different thresholds (95–99%) in comparing and selecting optimal thresholds for the four facial skin regions. Ideally, for

present purposes, a threshold should still be used to reduce the number of false negatives, as a skin-based biometric system should accommodate as many skin images as possible. Because a fixed threshold is difficult to set, 98% was chosen as a trade-off between system accuracy and population coverage for all facial skin regions in the XM2VTS database. Metric measures such as Precision, Recall, F-measure, FPR and FNR are detailed for the three schemes in Table 4.4. The results suggest that Scheme 3 is probably the best choice for assessment of ROI applicability for biometric purposes.

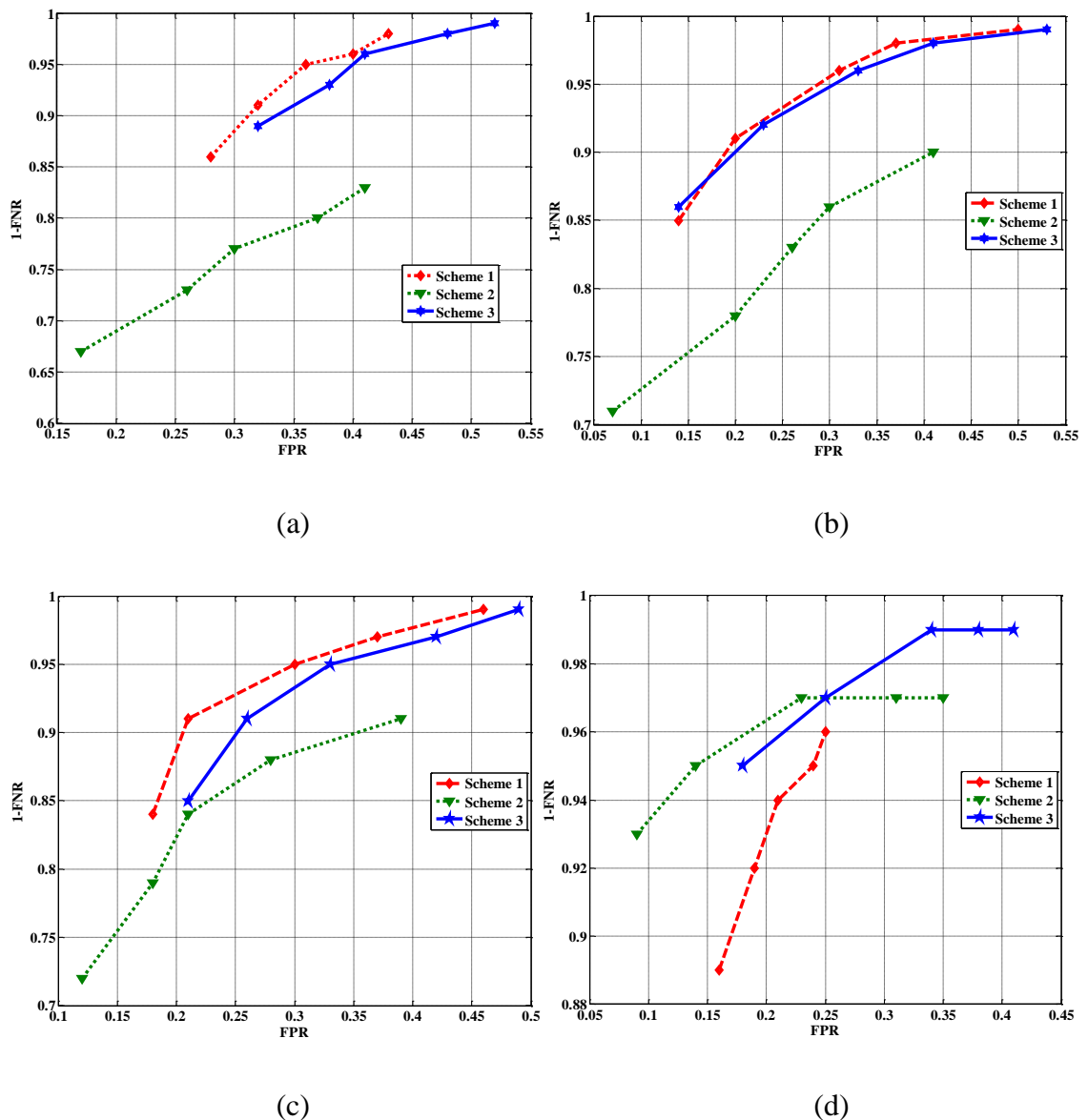


Figure 4.4 ROC curves of thresholding optimisation for (a) forehead, (b) left cheek (c) right cheek, (d) chin skin region using the three different schemes

ROC curves are commonly used to optimise the threshold parameter for the influence analysis. Figure 4.4 shows that although Scheme 1 performed better with forehead and cheek regions, it performed worse with chin regions. While Scheme 2 achieved better performance with chine, it performed generally worse than Scheme 1 and Scheme 3 with forehead and cheek regions. From these observations, scheme 3 seems to be an appropriate choice for skin purity assessment. In terms of optimising the thresholds, as five thresholds were tested (95-99%), FPR is still higher for all schemes. If the threshold is higher (97-99%), FPR is decreased indicating more accurate results. Consequently, 98% is chosen for skin purity assessment. These preliminary analyses justify the selection of the optimal value of the appropriate parameter for the proposed skin detection framework. The F-measure represents a balance between Precision and Recall, and the higher F-measure, the better the performance. Although precision is slightly higher for Schemes 1 and 2, Scheme 3 achieved the highest Recall for all four facial skin regions. Table 4.4 shows all these statistical measurements of the three investigated schemes.

Table 4.4 Skin detection framework performance of the three schemes for the four independent facial skin regions (the threshold value is set to **98%**)

ROI	Scheme	Precision	Recall	FPR	FNR	F-measure
Forehead	Scheme 1	0.83	0.91	0.32	0.09	0.87
	Scheme 2	0.83	0.73	0.26	0.27	0.78
	Scheme 3	0.81	0.93	0.38	0.07	0.87
Left Cheek	Scheme 1	0.99	0.91	0.20	0.09	0.95
	Scheme 2	0.98	0.78	0.20	0.22	0.87
	Scheme 3	0.98	0.92	0.23	0.08	0.95
Right Cheek	Scheme 1	0.99	0.90	0.21	0.10	0.94
	Scheme 2	0.99	0.79	0.18	0.21	0.88
	Scheme 3	0.99	0.91	0.25	0.09	0.95
Chin	Scheme 1	0.98	0.92	0.19	0.08	0.95
	Scheme 2	0.99	0.95	0.14	0.05	0.97
	Scheme 3	0.97	0.97	0.25	0.03	0.97

From the results presented in Table 4.4, it can be seen that Scheme 3 has achieved a better performance in terms of F-measure and Recall in comparison to other two schemes. It can also be observed that the performance of Scheme 3 of all four regions outperformed the other two schemes. This observation indicates that Scheme 3 is more applicable to deal with skin noise compared to Scheme 1 and Scheme 2 and it can be suggested that Scheme 3 is the best choice for assessing purity of facial skin region. The impact of Scheme 3 on population coverage will be also explained in Section 5.6.

4.7.3 Analysis of Computational Cost

In designing a system, it is important to consider in the process that how fast or slow particular algorithm performs. Any algorithm takes different time for the same input due to some factors such as processor speed, desk speed, compiler, etc. in this experiment, it is assumed that the machine used for the running the algorithm is affordable and the time complexity is computed by the execution time. The proposed framework was run in the MATLAB R2014a programming environment on a personal computer (operating system Windows 7 Enterprise, 3.20 GHZ, 16 GB memory 64-bit processor). The speed of the process was measured in seconds per ROI. Table 4.5 shows the computational cost for each of the three schemes.

Table 4.5 Computational time for the three skin detection schemes

ROI	Computation time (seconds)		
	Scheme 1	Scheme 2	Scheme 3
Forehead	0.132	0.122	0.147
Left Cheek	0.135	0.129	0.144
Right Cheek	0.134	0.133	0.139
Chin	0.144	0.125	0.153

4.8 Summary

The essential principles of skin pixel detection technique explained in this chapter have informed the proposed automatic framework for assessing ROI skin content for biometric purposes. The general skin pixel detection framework was detailed. Three

colour schemes were investigated. All three colour schemes were evaluated and validated using different databases. The technique for assessing skin purity of ROI was also described. Subjective observations and the mechanism used to generate the ground truth for purity analysis were also detailed. Finally, computational cost was also analysed.

In terms of the proposed skin pixel detection scheme (Scheme 2), the scheme's performance shows promise in skin pixel detection technique compared to the state-of-the-art methods.

Comprehensive experimental analysis was also made for each individual scheme. The analysis was covered all three colour schemes for skin purity assessment. Skin regional purity's performance is analysed in terms of Recall and Precision. In general, the proposed technique promises to enhance human interaction and is likely to be more accurate. Although all the colour schemes returned comparable results, Scheme 3 proved most effective for skin purity assessment and is preferable in terms of population coverage and as a trade-off among alternative schemes. To conclude, the proposed technique (Scheme 2) can determine automatically regions excessively corrupted by non-skin pixels, which are excluded from biometric processing. In the next chapter, the proposed technique will be implemented in a skin-based biometric system.

Chapter 5

Exploring the Potential of Facial Skin Regions for Identity Information

5.1 Introduction

This chapter describes the design, implementation and evaluation of a prototype skin-based biometric identification system to investigate the use of textural information from facial skin regions as a source of biometric information. This information is likely to be of interest when the entire facial image is not available, as for instance in forensic investigations or surveillance images. Four regions were chosen to obtain facial skin texture: forehead (R1), right cheek (R2), left cheek (R3), and chin (R4), as these are relatively easy to distinguish and at least some of these are likely to be visible even in partially occluded facial images. To facilitate extraction of these regions of interest, facial landmarks are automatically detected. A skin purity assessment technique, proposed in Chapter 4, is applied to identify the regions with sufficient skin content. Each of the available skin regions is analysed independently in attempting to establish the individual's identity, using features based on Local Binary Patterns (LBP) and Gabor wavelet filters. Fusion of features from the skin sub-regions is applied prior to classification of the images.

Experiments were performed using the publicly available the XM2VTS database to evaluate the performance of the skin-based biometric recognition system. As the proposed skin detection technique yielded promising and comparable results to other state-of-the-art techniques, this method is incorporated with the skin-based biometric system for people identification proposed in this chapter. Of the regions investigated, the forehead and chin were found to provide the richest biometric information.

Facial recognition (FR) has many practical applications and is a well-established research problem in computer vision. The most common FR technique is based on holistic facial features, but in practice, it can be difficult to obtain the whole facial image, which is often occluded by objects, accessories (e.g., sunglasses, hat, scarf) or facial hair [175]. Occlusion caused by these issues may distort or hide some characteristics of facial images. At this point, a system based solely on textural information extracted from local features and dealing with part of the face would be of considerable value in identifying individuals from partial images, using an automated facial skin-based recognition system. Matching two facial skin regions is very challenging because it requires the extraction of distinctive features from skin images that are uncalibrated because of the variations in illumination, posture, facial hair, and accessories. Another challenge for skin biometric systems is that intra-class variations are still larger than inter-class variations [176]. An ideal skin feature representation should be capable of discriminating between different individuals. While researchers commonly use pure skin regions and avoid any facial skin noise, this is likely to be impractical in real-world applications [5] [21] [175]. Therefore, establishing systems to automatically deal with these issues are expected to standardise the examination process and also make the process more efficient.

For texture analysis, skin image resolution is crucial, as this determines the level of skin detail that can be extracted from the image. A common scientific measure of face image resolution is *inter-pupillary distance*: the distance between the centres of the two pupils. Face images are described as high resolution when this distance is greater than 60 pixels [39]. ISO/ICAO standards also recommend a minimum inter-pupillary distance of 90 pixels for facial images in travel documents [51]. Li et al. [7] proposed a system for defining skin texture in terms of pore-scale facial features, extracted from pores, fine

wrinkles, and hair using PCASIFT. They investigated the hairless cheek region using a number of databases such as Multi-PIE [38], which contains high-resolution images with an inter-pupillary distance exceeding 400 pixels. As a result of these considerations, the XM2VTS database [171] has been used in the present study which contains facial images with inter-pupillary distances in the range 84–126 pixels.

The contribution of this chapter can be summarised as follows: (1) An automated skin-based biometric system is proposed for use when the whole face image is not available, and (2) the system automatically segments the regions of interest and investigates the possibility of using each region as a source of biometric information in cases of partial occlusion by exploring likely visible skin regions.

The chapter is organised as follows. Section 5.2 describes a general framework for the skin-based biometric system, including detection of facial landmarks, localisation of region of interest, feature extraction, classification, and evaluation and validation of the skin-based biometric system. Section 5.3 presents the experimental setup. The results obtained from both skin pixel detection framework and biometric system performance are then analysed in Section 5.4. Further improvement of biometric system and results of the analysis are described in Section 5.5 and Section 5.6 respectively. Finally, conclusions of this chapter are drawn in Section 5.7.

5.2 General Framework of the Proposed System

The schematic of the proposed system shown in Figure 5.1 indicates all the stages of the framework, each of which is subsequently described in detail. The framework comprises four principal stages: (i) pre-processing, (ii) skin purity assessment, (iii) feature extraction and image representation, and (iv) classification. In the pre-processing stage, the facial image is captured, and the facial landmarks are detected to automatically localise the ROI based on facial geometric measurements. In the second stage, skin purity analysis assesses whether the input includes sufficient skin for biometric processing. The third and fourth stages are feature extraction and classification respectively which are widely used in biometric systems. In the feature extraction stage, the skin image ROI is further split into several non-overlapping sub regions to capture the skin's textural detail.

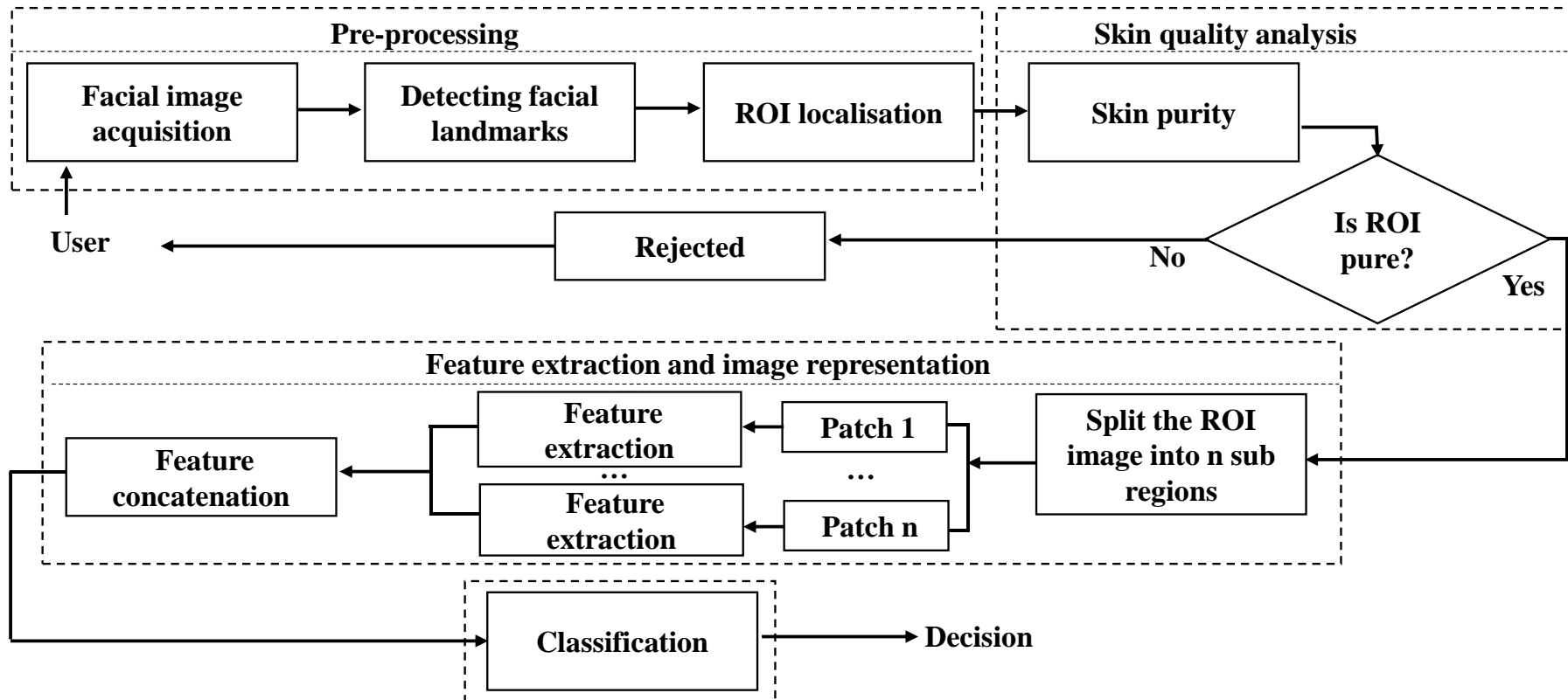


Figure 5.1 Framework of the proposed skin-based biometric system

The skin feature is then extracted from each sub region, all of which are concatenated to a final single feature vector. Finally, the classification stage classifies these features for person identification.

5.2.1 Detecting Facial Landmarks for ROI Segmentation

Automatic extraction of facial landmarks is a fundamental step in designing a fully automated recognition system. This process can readily support real-world applications such as face recognition, criminal identification, and surveillance. In many such applications in real scenarios, facial feature detection and localisation are crucial, and some competing techniques seem to be more effective for automatic facial feature learning. Among these, Active Shape Model (ASM) [18] and Active Appearance Model (AAM) [19] are widely used to detect facial landmarks. AAM analyses texture information using pixel intensity within the face region while ASM uses the shape information from the entire facial image. Most recently, Chehra Face Tracker [20] [177] has emerged as an efficient method of localisation using cascade linear regression to learn the mapping, using both texture and shape information from the facial image. Chehra works with frontal facial images and provides 49 different facial landmarks and 10 eye landmark points. In the proposed scheme, extracted facial landmarks using Chehra software are used to correctly locate the region of interest (ROI). In this context, ROIs are chosen to be those areas that contain the most visible skin regions on the face—the forehead, right cheek, left cheek, and chin region. Facial landmarks detected by Chehra, such as eye corners, nasal bridge, and mouth corners are utilised to precisely locate these four ROI regions. Figure 5.2 shows some examples of facial landmarks detected by the software.

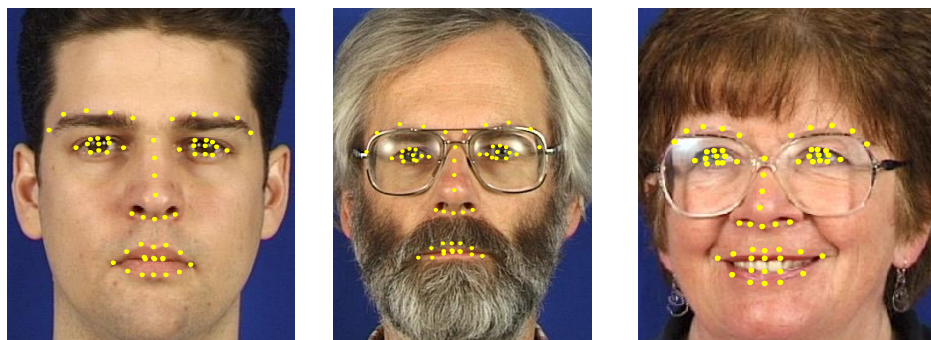


Figure 5.2 Facial landmarks detected by Chehra Face Tracker

5.2.2 Geometric Skin Localisation Measurements

It is worth noting that the preliminary processing step uses the coordinates of both eyes to normalise the facial images. As a first step, all facial images are automatically aligned horizontally, using the eye-centres as reference points. Geometric measurements for ROIs are based on the geometric model proposed in [178]. Figure 5.3 depicts the geometric measurements for segmenting each region within the facial image. Assuming that the detected images are frontal-view faces, the selected regions will be determined as follows.

Forehead region (R1): Let (x_{reye}, y_{reye}) and (x_{leye}, y_{leye}) be the outer corner of each eye (right and left respectively), and let $d1$ be the distance between them. This distance was computed as:

$$d1 = [(x_{reye} - x_{leye})^2 - (y_{reye} - y_{leye})^2]^{0.5} \quad (5.1)$$

The vertical distance between the outer corners of the eyes and the rectangular forehead region (R1) is set to $0.3d1$. The size of the rectangular forehead region, which is located exactly $0.3d1$ above the eye corners, is $d1 \times 0.4d1$ pixels.

Right (R2) and Left (R3) cheek regions: Let (x_{upper}, y_{upper}) and (x_{lower}, y_{lower}) be the two selected points on the nasal bridge (dorsum nasi). The distance between these points is calculated as:

$$d2 = [(x_{upper} - x_{lower})^2 - (y_{upper} - y_{lower})^2]^{0.5} \quad (5.2)$$

The locations of right cheek (R2) and left cheek (R3) regions are set at $d2$ from the nasal bridge respectively. Each region will have dimensions of $d2 \times d2$ pixels.

Chin region (R4): Let (x_{rm}, y_{rm}) and (x_{lm}, y_{lm}) be the two corner points of the mouth region. The distance between the mouth corners is calculated as:

$$d3 = [(x_{rm} - x_{lm})^2 - (y_{rm} - y_{lm})^2]^{0.5} \quad (5.3)$$

The size of the chin region (R4) is $0.55d3 \times 0.35d3$ pixels, located at a vertical distance of $0.35d3$ beneath the mouth region.

ROI size and location parameters are optimised empirically to ensure that they fall within the facial boundaries of all faces in the database. After detecting the facial landmarks, the four regions (R1–R4) are cropped for further analysis.

Table 5.1 shows all the parameters of the four regions, including their dimensions and locations.

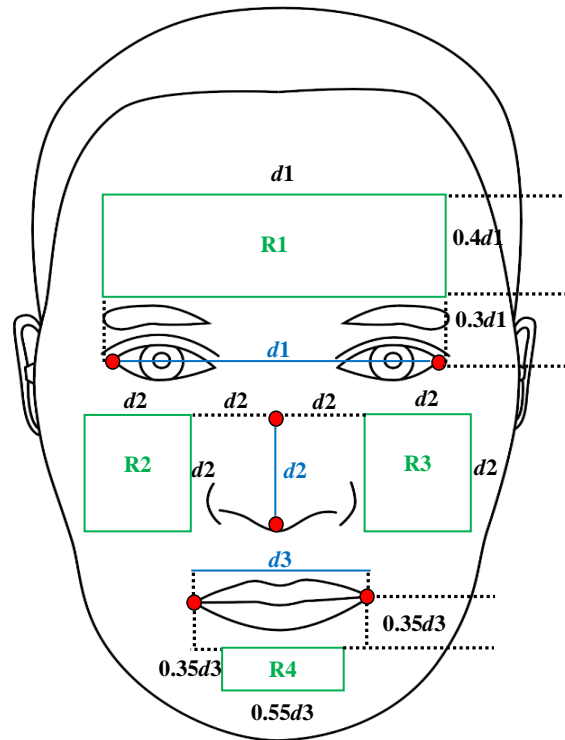


Figure 5.3 Geometric facial landmark measurements for ROI localisation. These measurements were empirically determined using Chehra software. All ROIs fall within the face boundary

Table 5.1 ROI dimensions based on different face landmarks, where $d1$ is the distance between the two eyes, $d2$ is the length of the dorsum nasi, and $d3$ is the distance between the outer mouth corners

ROI	Key facial landmarks	Dimensions(pixels)	Location
Forehead	outer corners of two eyes	$d1 \times 0.4d1$	$0.3d1$ above eye centres
Cheeks	nasal bridge (dorsum nasi)	$d2 \times d2$	$d2$ away from nasal bridge
Chin	corners of the mouth	$0.55d3 \times 0.35d3$	$0.35d3$ below the corners of the mouth

5.2.3 Skin Pixel Detection for Skin Purity Analysis

This phase investigates the suitability of the extracted ROI for further processing. As explained in Section 4.5, the proposed skin pixel detection technique (termed Scheme 2) employs two colour spaces: HSV and YCbCr in this study. In real-world applications, choosing a single colour space to detect skin pixel regions may not overcome all the challenges of skin pixel detection. Applying more than one colour space not only helps to eliminate some of these issues but also improves detection performance. To identify the skin pixels, RGB values for each pixel are converted to HSV and YCbCr values. Once these values have been computed, each pixel is classified as skin or non-skin by comparing Cr, Cb, and H values against a set of empirically determined thresholds. Morphological operations (i.e. *opening* followed by *majority* operation, using a 3x3 square structuring element) were also applied to eliminate incorrectly marked pixels within the skin region. These operations help in removing or filling small groups of pixels caused by noise within the facial image, such as a strand of hair or facial marks (e.g. freckle, mole, scar).

As the skin purity assessment takes place, if the proportion of skin pixels in the ROI is less than the purity assessment threshold, those regions are then excluded from subsequent biometric analysis.

5.2.4 Feature Extraction

If the ROI is classified as a pure skin, the feature extraction is applied after converting the image to greyscale. Irrespective of the ROI size, each is divided into four sub-regions, each of which is treated separately for feature extraction. The generated feature vectors are concatenated to form a single feature vector. The two techniques utilised for feature extraction are Local Binary Pattern (LBP) and the Gabor wavelet filters.

5.2.4.1 Local Binary Pattern (LBP)

The advantage of the LBP method is that it is a powerful descriptor for extracting local texture features. The original LBP operator labels each pixel by thresholding it within a 3x3 window, using the centre pixel value to obtain the result as a binary number. The

resulting binary numbers for all pixels are converted to corresponding decimal numbers, which can be used to build a histogram for texture description [88]. LBP was subsequently developed to accept any window size [113]. The descriptor was developed to define local neighbours as a set of sampling points (P), evenly spaced on a circle centred at the pixel to allow any radius (R) and any number of points. If the point does not fall in the centre of a pixel, this is resolved by means of bilinear interpolation.

Formally, given a pixel at (x_c, y_c) , the resulting LBP is formulated as [88] [113]:

$$LBP(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \quad (5.4)$$

where g_c, g_p correspond to the grey-scale of the central pixel $c(x_c, y_c)$ and the P neighbourhoods pixels, respectively. The function $s(x)$ is defined as [113] :

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \quad (5.5)$$

5.2.4.2 Gabor Filters

Gabor functions, which have been successfully employed in biometric systems, are mathematically attractive for minimizing the joint space-spatial frequency uncertainty [118] [179] [180]. A two-dimensional Gabor filter is derived from a Gaussian kernel function modulated by a complex sinusoidal plane. The feature vector is constructed from the concatenation of all the means and the standard deviation values of Gabor bank filter response magnitudes. This technique was used in [181] and proved robust for extraction of texture features. Formally, a 2D Gabor filter ψ in the spatial domain can be expressed as:

$$\psi(x, y; f, \theta) = \frac{f^2}{\pi\gamma\eta} e^{-(\alpha^2 x' + \beta^2 y')} e^{j2\pi f x'}$$

$$x' = x \cos\theta + y \sin\theta \quad (5.6)$$

$$y' = x \sin\theta + y \cos\theta$$

where f is the central frequency of the sinusoidal plane wave, θ represents the orientation angle of the Gaussian major axis with respect to the plane wave, α and β are the sharpness along the major axis, and the sharpness along the minor axis (perpendicular to the wave) respectively. In the given form, the aspect ratio between the frequency and sharpness is constant. $\eta=f/\beta$ and $\gamma=f/\alpha$. The skin image ζ , $\zeta(x,y)$ can be convolved with the 2-D Gabor filter function using a bank of scale u and orientation v . $f_u = f_{max} (\sqrt{2})^{-u}$ and $\theta = \frac{k\pi}{n}$ where $k = \{0, \dots, n-1\}$. Therefore, the convolution can be expressed as:

$$r_\zeta(x,y; f_u, \theta_v) = \psi(x, y; f, \theta) * \zeta(x, y) \quad (5.7)$$

All means μ_{uv} and standard deviations σ_{uv} of the magnitude $|r_\zeta|$ are computed to construct the skin feature vector F .

5.2.4.3 Feature Representation for LBP and Gabor Filters

The dimension of the feature vector depends on the method used for feature extraction. For example, for the LBP operator with parameters $P = 8$ and $R = 5$, the vector dimension is $2^8 = 256$ for each sub-region. Following feature-level fusion of the four sub-regions, the dimension becomes 1024 ($= 256 \times 4$). Gabor wavelets of eight different scales $\{u \in 0, \dots, 7\}$ and eight orientations $\{v \in 0, \dots, 7\}$ were used, resulting in a feature dimension of 128 ($= 8 \times 8 \times 2$). Other parameters were set as $f_{max} = 0.25$, $\gamma = \sqrt{2}$, $\eta = \sqrt{2}$. The feature vector was then constructed from the concatenation of the four sub-regions, generating a 512-dimension feature vector. This technique was previously used in [182] and proved robust for extraction of texture features.

5.2.5 Skin Features Classification

Two types of classifier were used here for biometric identification by using skin texture: k-Nearest Neighbour (k -NN) classifier ($k = 1$), and Support Vector Machine (SVM) [12]. Optimal kernel selection for SVM is performed by testing different kernels, such as sigmoid, polynomial, Gaussian (Radial Basis Function), and linear kernels. Following trials, the polynomial function was selected as optimal.

5.3 Experimental Setup

In order to analyse the performance of skin biometric system, the XM2VTS [171] database facial images are utilised in this work. As already described in Section 4.6, the XM2VTS is a multi-modal database comprising still colour images, video sequences, and speech recordings from 295 individuals. The total number of the frontal images is 1,180 (295 individuals, 4 images each).

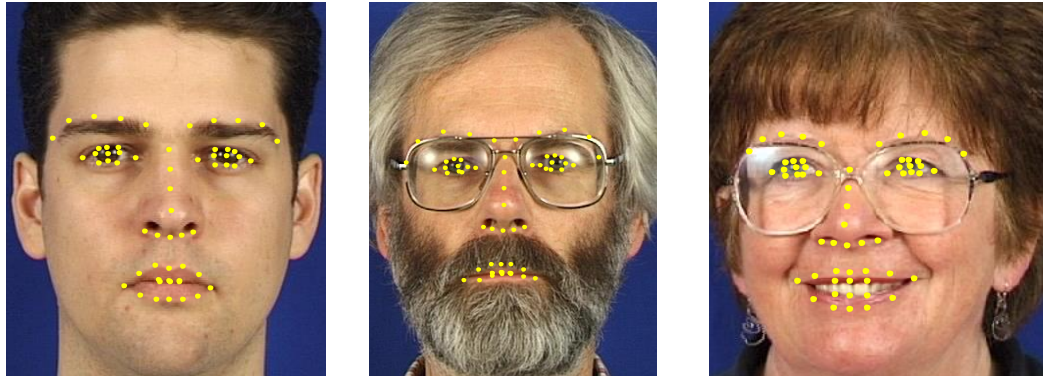
Chehra Face Tracker software [20] [177] can detect facial landmarks in only 1,128 of the 1,180 available images. All four images were marked for only 274 of 295 individuals in the XM2VTS database. These 274 individuals were used for the experiments in this thesis and the remaining 21 were excluded. As explained in the previous section (Sub-sections 5.2.1 and 5.2.2), the four ROIs were segmented from the face images, and these regions may differ in terms of noise such as hair or facial marks (e.g. moles, wrinkles, scars). As the distance between the two eyes ranges between approximately 84 and 126 pixels, the size of extracted facial skin regions may also vary between images.

5.4 Experimental Analysis

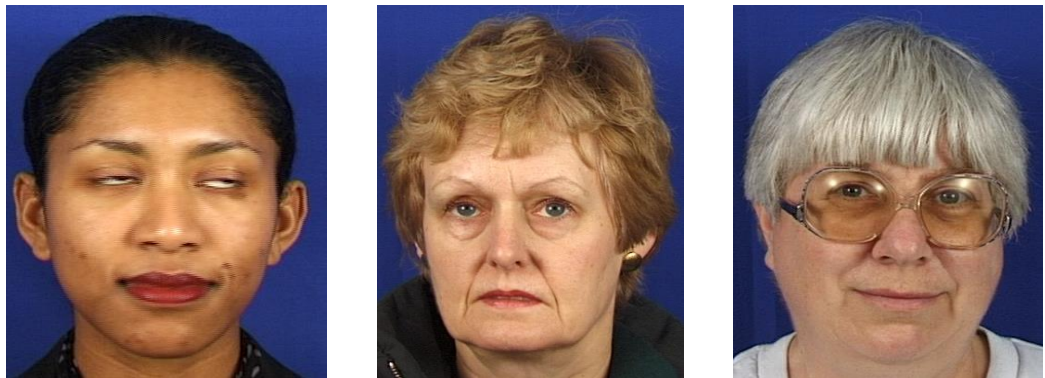
As outlined above, the main objective of the present study is to construct a facial skin-based biometric system, both to overcome the limitations of facial recognition systems in cases such as occluded view or noise in the facial area and to supplement existing facial recognition systems by providing additional information.

5.4.1 Facial Landmark Detection

Figure 5.4 shows some examples of users who were detected and not detected by Chehra Face Tracker software. It was noticed that blurring, reflections, and head pose impacted negatively on the effectiveness of Chehra Face Tracker software [20] [177]. The exclusion of 21 individuals was an error termed as failure to acquire (FTA), where the landmarks in a face image cannot be detected. Using Chehra software, the FTA rate was 4.40 % that is 52 of 1,180 images were not detected.



(a) Facial landmarks were detected



(b) Facial landmarks were not detected

Figure 5.4 Some examples of users (a) whose facial landmarks were detected and (b) were not detected by Chehra Face Tracker software

5.4.2 Skin Purity Analysis

As a preliminary step, first, only the proposed skin pixel detection scheme (Scheme 2) was applied. The skin pixel detection scheme output is returned in the form of a binary mask (1 = skin pixels, 0 = non-skin pixels). Figure 5.5 shows the algorithmic flowchart for this purity analysis process.

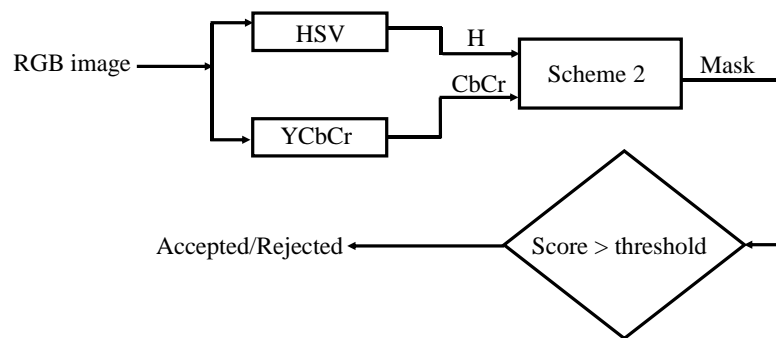


Figure 5.5 Flowchart of the proposed purity assessment technique using skin pixel detection Scheme 2

ROI images containing non-skin pixels greater than a defined threshold were rejected from subsequent analysis. Users will be required to provide a new sample in such cases. All the accepted images were used to assess the performance of the proposed biometric system. Figure 5.6 shows samples of the four regions with differing levels of noise and indicates whether these were rejected or accepted for further processing.

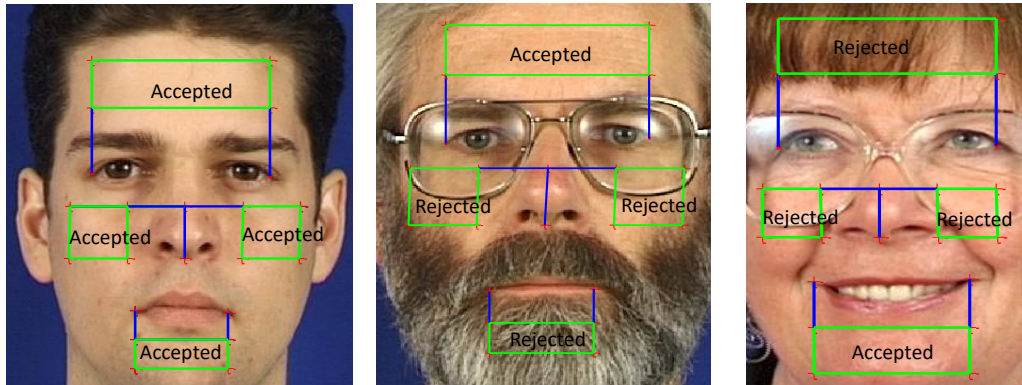


Figure 5.6 Examples of accepted and rejected ROIs based on the skin purity assessment technique

The forehead region is less likely to be affected by facial expressions while cheek regions are prone to noise (e.g. sporadic marks, facial hair, shadows, or other artefacts) that could impact negatively on performance.

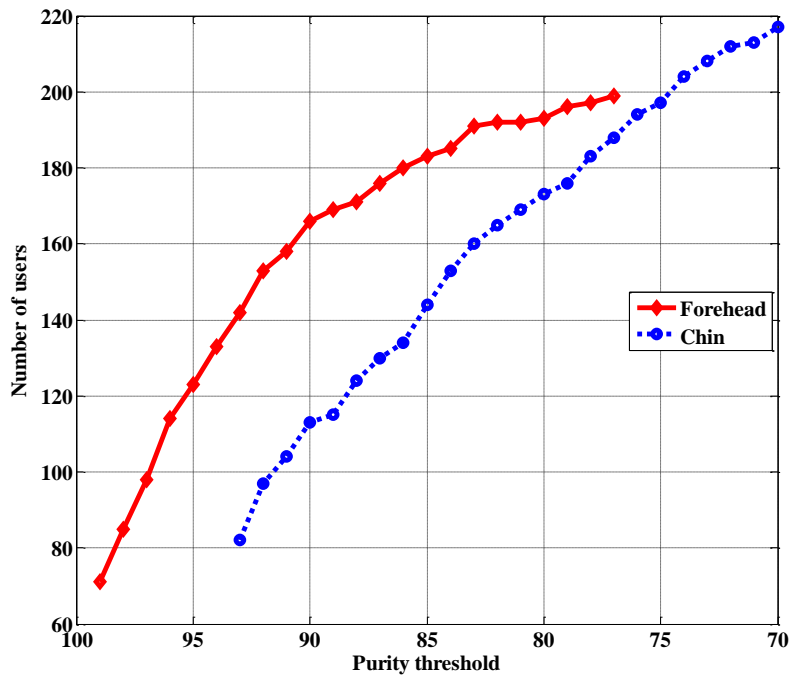


Figure 5.7 Subjects with useable regions by set of thresholds for forehead and chin region

In males, the chin region contains hair roots that make the skin region darker, and the skin detector is likely to classify some of these pixels as non-skin. In this analysis, different selective threshold values were used to control acceptable non-skin corruption in the extracted ROIs. Figure 5.7 shows an example of two different skin regions (forehead and chin) and the percentage of skin pixels in a specific ROI for all images of these individuals. The data suggested that different skin purity thresholds may be appropriate for different regions when using Scheme 2 for skin pixel detection technique. Experimental results for this skin pixel detection technique were deeply explained in Section 4.7.

5.4.3 Performance of Skin Biometrics

All images accepted as pure skin are then input to the biometric system for person identification. The optimised skin purity assessment technique is incorporated into the proposed skin-based biometric system and tested using facial images from the XM2VTS database. The experimental protocol uses three images from each person as gallery (training) set; the remaining images are used as a probe (testing) set. Where the ROI is recognised as pure skin, the region is split into four sub-regions for feature extraction, irrespective of ROI size.

Table 5.2 presents the rank-one recognition rates, with an example of the chosen threshold for skin purity assessment technique for each of the four facial skin regions. For instance, a purity threshold of 90% skin pixels was selected experimentally for the chin region while the threshold for the forehead region was also set experimentally at 97%. The reason for selecting different threshold values is that the number of users included in the evaluation for the four ROIs was approximately similar. Note that the similarity of population coverage makes the comparison between the four facial skin regions for the skin biometric performance to be reasonably fair. The algorithm was run several times, and average recognition rates are reported.

The results suggest that classification accuracy is much higher for the forehead and the chin regions than those for the cheek regions. The highest rate of recognition accuracy (91.33%) was for the forehead region. The chin region ranked second for recognition accuracy (80.97%) when using the LBP descriptor features with SVM classifier. Both

cheek regions returned lower accuracy rates and this possibly because the cheeks are more affected by variations in lighting, facial expression, and posture, presence or absence of spectacles, and so on. This means that the forehead and chin are likely to provide more robust biometric information in cases where the face is only partly visible, and these were therefore selected for further exploration.

It has been observed that when compared to SVM, the Gabor filters did not perform well on the nearest neighbour classifier (k -NN), and this aligns with the results reported in [183] and [181].

Table 5.2 Recognition accuracy (rank-one) at a defined threshold for the four ROIs (%)

ROI	Skin purity Threshold (%)	No. of users	LBP		Gabor Wavelet	
			k -NN	SVM	k -NN	SVM
Forehead	97	98	75.77	91.33	23.70	66.58
Right Cheek	100	104	44.47	60.34	19.71	54.33
Left Cheek	100	89	41.01	57.58	23.03	46.91
Chin	90	113	60.62	80.97	26.33	74.12

The LBP performed much better than Gabor wavelet for skin texture information extraction, and SVM also performed better than the k -NN classifier. As they provided the best accuracies, Figure 5.8 shows rank- n recognition rates for the two ROIs (forehead, and chin) based on LBP operator using k -NN and SVM classifiers. The results confirm the superior performance of the forehead region for individual's identification.

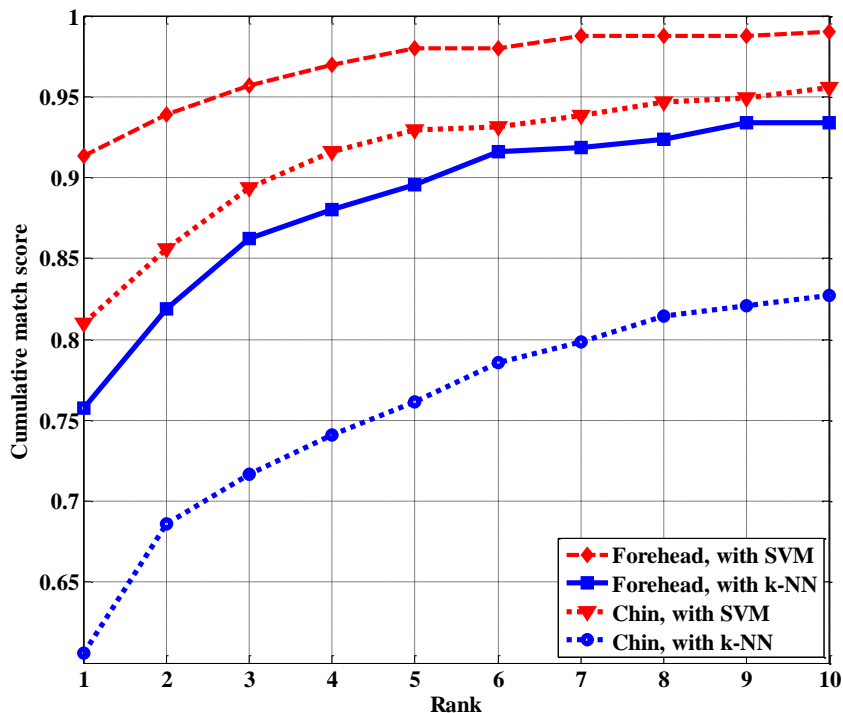


Figure 5.8 CMC curves of recognition rates for the two ROIs (forehead and chin) based on LBP operator using k -NN and SVM classifiers

Further investigation using different thresholds of skin purity with the forehead and the chin region (see Figure 5.9) suggested that recognition rates are affected by the use of lower skin purity threshold. Increasing the purity skin threshold, the identification performance was improved leading to higher accuracy for biometric system.

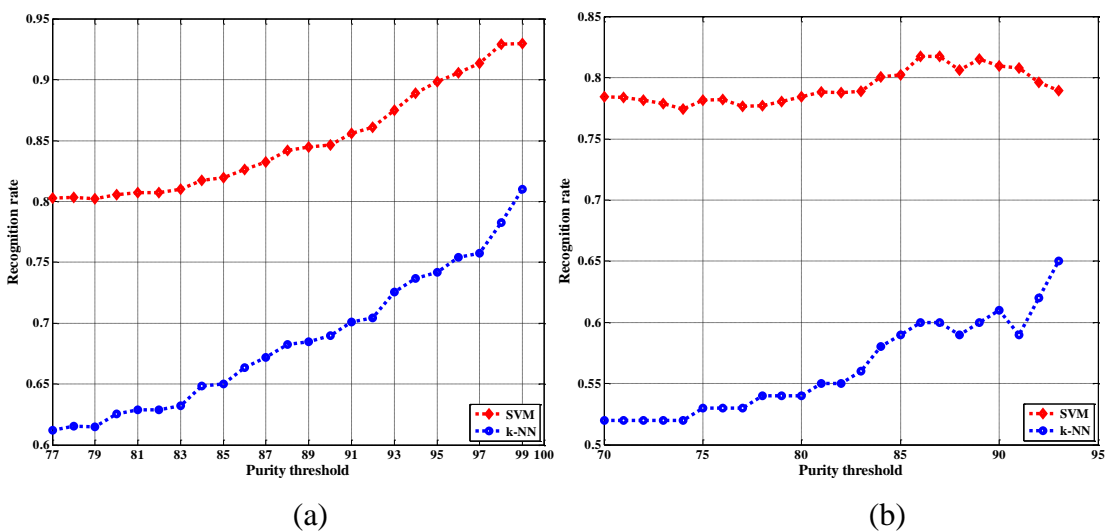


Figure 5.9 Recognition rates for (a) forehead and (b) chin region based on LBP operator using two different classifiers (k -NN and SVM) with different thresholds

5.4.4 Comparison with Existing Systems

The proposed system achieves better performance than previously published alternatives. Table 5.3 compares the proposed system framework with earlier work. For fair comparison, only the results of the proposed scheme (Scheme 2) based on LBP for feature extraction and SVM for classification are compared. The comparison shows that the cheek regions (right and left) provided slightly lower but comparable accuracy (60.34% and 57.58%, respectively).

Although the chin region has not (to the best of the author's knowledge) been investigated before, it has delivered better performance within the proposed system framework, confirming its utility as a source of biometric information. When compared to previously published results, recognition rate for the forehead region is significantly higher (by about 23%), indicating the superiority of the proposed technique.

Table 5.3 Proposed system's recognition performance compared to earlier methods using the same database (XM2VTS)

Algorithm	System Parameters		Skin ROI	Accuracy
	Features method	Classification method		
Lin et al. [5]	Regularised LDA	Adaptive metric fusion	Cheeks + Forehead	0.67
Al-Qarni et al. [184]	Gabor filter	k -NN	Forehead	0.68
<i>Proposed system</i>	LBP operator	SVM	<i>Forehead</i>	<i>0.91</i>
			Right cheek	0.60
			Left cheek	0.58
			<i>Chin</i>	<i>0.81</i>

5.5 Further Improvement using Multi Scheme Purity Assessment for Skin-based Biometrics

The experimental results presented in Section 5.4 demonstrated that the skin-based biometric system is capable of distinguishing individuals. The main objective of these experiments presented in this section is to incorporate more colour spaces in skin purity

assessment technique to overcome the failure-to-acquire errors that occur when only scheme 2 for skin pixel detection was used. This is achieved by introducing an enhanced skin pixel detection technique (Scheme 3). This scheme was explained in Section 4.5 and its performance was analysed in Section 4.7. Additionally, a common skin purity assessment threshold for all ROIs is also explored. It is clear that setting a common threshold may be a better solution to avoid the selection of a threshold which might be data dependant. All these concerns were tackled in this section with further investigation.

5.5.1 Improving the Skin Purity Analysis Assessment

The skin purity assessment technique assesses whether the extracted ROI is unduly occluded by artefacts, either natural (e.g. hair, moles) or artificial (e.g. headgear). Unlike other studies, in which features like moles and birthmarks were used as the discriminatory biometric features, the present work explores the suitability of pure skin region as a source of biometric information, since moles and other distinguishing marks are not universal across individuals. Note that such artefacts can be imitated easily. In this implementation, skin colour models were used to ascertain the level of purity of the extracted skin region of interest. While an individual's skin tone typically occupies a compact region in the colour space, skin appearance can be altered by factors such as illumination or camera sensor, as well as by the individual's age, sex, ethnicity, and so on. As skin colour is largely invariant to partial occlusion, scaling, pose, and rotation, it permits more reliable differentiation between skin pixels and other artefacts within a facial image. Other sources of variability (e.g. exposure to weather, especially sunlight) are not considered here, as all of the database images used were captured in a controlled environment.

Given the evidence that use of more than one colour space helps to eliminate some of the above concerns and so improve skin pixel detection performance, the enhanced skin pixel detection technique (i.e. termed Scheme 3) presented here uses three colour spaces: RGB, HSV and YCbCr. Although the RGB colour space (Scheme 1) is widely used in image segmentation, one of its limitations is that luminance and chrominance information are not separated. However, HSV and YCbCr (Scheme 2) contain pure

colour information and should therefore facilitate separation of skin pixels from artefacts for the purposes of colour-based skin purity assessment analysis. Therefore, the two schemes generated a binary mask in which morphological operations (i.e. *opening* followed by *majority* operators, using a 3x3 square structuring element) were applied to eliminate small groups of pixels that were incorrectly marked (because of noise or small dots such as moles) within the skin ROI. In Scheme 3, these two masks were combined to take a final decision. Scheme 3 was then applied to assess whether the skin in the ROI was adequate for subsequent processing (see Sections 4.5 and 4.7 for more detail).

5.6 Experimental Results and Discussion

This section presents the results obtained from the proposed skin-based biometric system using Scheme 3 for skin purity assessment. Results of skin purity assessment technique are discussed. Biometric performance is also analysed. Finally, a comparative study is presented.

5.6.1 Skin Region Purity Analysis

A combination of two colour schemes (Scheme 3) was used to determine the purity of skin regions extracted from the facial images. Where the proportion of skin pixels was higher than a chosen threshold, the ROI was deemed to be of adequate quality (accepted) for subsequent processing. Performance was measured using True Positive Rate (TPR) and True Negative Rate (TNR). True positive (TP) refers to cases where the colour models correctly classified an acceptable purity ROI, and true negative (TN) refers to cases where a non-pure region was correctly classified (rejected). Additionally, Half Total Error Rates ($HTER = (FPR+FNR)/2$) were also computed to determine the optimal threshold for skin purity assessment technique.

Taking two different facial skin regions as examples (i.e. forehead and chin region), Figure 5.10 and Figure 5.11 show TPR and TNR for all three colour schemes separately. It is observed experimentally that, while the scheme1 performed better in identifying pure forehead skin regions, the Scheme 2 performed better for occluded regions. It is also found that Scheme 2 performed better than Scheme 1 in identifying

pure chin skin region. As expected, TPR fell gradually and TNR increased as the threshold was increased. Additionally, from the results shown in Figures 5.10 and 5.11, it can be seen that, for TPR, Scheme 1 and Scheme 2 do not always lead to an improved performance. It is clear that, for both regions, Scheme 3 achieved the top performance in comparison to the other two schemes. This observation indicates that more skin images can be recognised by using Scheme 3 for skin purity assessment. In terms of TNR, Scheme 2 performed relatively better than the other two schemes. Considering these measures and to mitigate the influence of using either Scheme 1 or Scheme 2, Scheme 3 is the best choice as a trade-off for skin purity assessment.

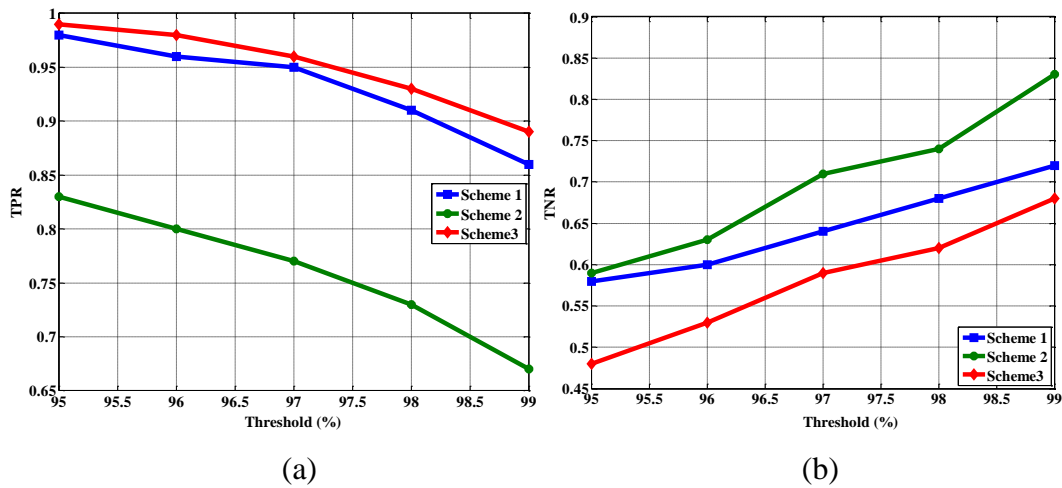


Figure 5.10 (a) TPR and (b) TNR of purity assessment schemes for various threshold values of forehead region

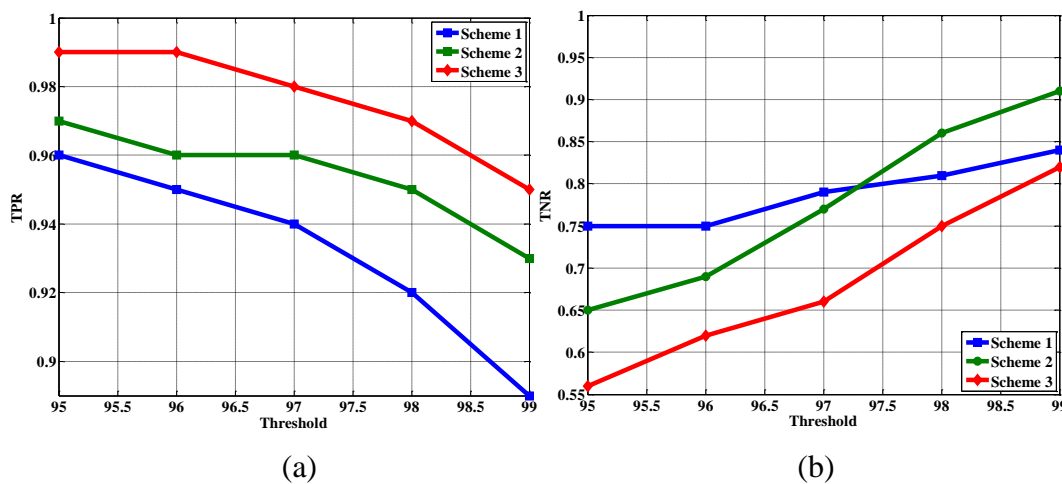


Figure 5.11 (a) TPR and (b) TNR of purity assessment schemes for various threshold values of chin region

In conclusion, observations from the results presented in Figures 5.10 and 5.11 suggested that according to the performance analysis measures applied in this experiment include TPR and TNR, Scheme 3 always achieved the highest performance for identifying skin pixels (TPR) with different facial skin regions. However, Scheme 1 and Scheme 2 are less stable which leads to misclassification with different facial skin regions. Scheme 2 is likely more robust for identifying non-skin pixels (TNR) compared the other two schemes. This observation can be seen clearly in Figure 5.10 (b) and Figure 5.11 (b) when the purity threshold is above 97%. Using Scheme 2, the more users will be rejected to be processed for biometric identification. In this case, Scheme 3 will mitigate the limitation of Scheme 1 and Scheme 2. Therefore, Scheme 3 is selected as a proper choice for skin purity assessments.

To identify the ideal threshold, HTER values of all four facial skin regions were plotted for different threshold settings (see Figure 5.12). Unlike TPR and TNR, which are measures of success, HTER values provide a composite measure of false positive (FP) and false negative (FN) errors. As Scheme 3 was chosen, HTER for this scheme has been investigated to analyse the optimal value of the threshold for skin purity assessment for the four facial skin regions. By comparing HTER between these regions, it can be seen that the error on the forehead region was higher than the other regions. The lowest error rate achieved by chin region. HTER rates initially dropped quite rapidly as the threshold increased but remained relatively stable once the threshold reached 97% or greater.

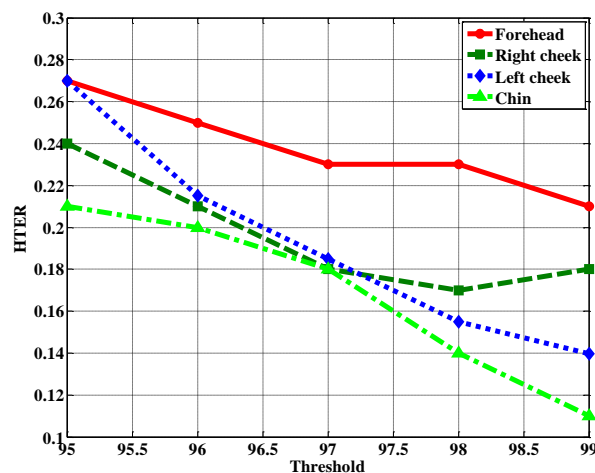


Figure 5.12 HTER of purity assessment scheme 3 for various threshold values of facial skin regions

To conclude, it can be seen from the Figure 5.12 that, by comparing HTER for all four regions, forehead region has generated higher errors compared to other three regions. Chin region has achieved the lowest error with all thresholds. Generally, the error was more stable if the threshold was above 97%. Therefore, 98% has been chosen for skin purity threshold.

In practice, a biometric system should cover across large population. Using Scheme 3, therefore, the population coverage is dramatically increased as skin biometric system can work in more robust conditions. Table 5.4 shows the number of individuals obtained from the three schemes indicating the advantages of using Scheme3 in skin biometric system.

Table 5.4 Population coverage of the three skin purity assessment schemes

Purity scheme	No. of individuals (purity threshold = 98%)			
	Forehead	Right cheek	Left cheek	Chin
Scheme 1	155	203	198	203
Scheme 2	104	158	145	222
Scheme 3	167	206	201	229

As all these numbers of individuals identified in the database as having 4 adequate ROI images, Scheme 3 improved the population coverage compared to the other two schemes. Individuals identified by Scheme 3 could be the optimal choice to investigate the stability and robustness of the proposed skin-based biometric system.

5.6.2 Biometric Performance Analysis

As demonstrated in Table 5.4, all individuals identified by the purity assessment schemes as having 4 ROI regions of acceptable purity were then inputted to the biometric system. The threshold of acceptable skin regions is derived from the statistical measures (HTER) and is set to 98% which is expected to be more noise-free. Identification task has been investigated in this set of experiment. Of the 4 available images for each individual, 3 images were used for enrolment, and the fourth was used for testing. The process was repeated several times; average recognition rates are reported in Tables 5.5, 5.6, 5.7, and 5.8. The process used LBP and Gabor wavelet

features. The LBP feature vector had a dimension of 1024, derived from four sub-regions (256×4) and the Gabor wavelets had a dimension of 512 (128×4).

Three classifiers has been used in this set of experiments for biometric identification based on skin texture: k -Nearest Neighbour (k -NN) ($k = 1$); Sparse Representation Classifier (SRC); and Support Vector Machine (SVM). By testing different kernels, the polynomial function was selected as the optimal kernel for SVM.

The experiments investigate the effect of all three skin purity assessment schemes. Considering the results shown in Tables 5.5, 5.6, 5.7, and 5.8, it is found that while Scheme 1 and scheme 2 achieved relatively better performance, they enrolled fewer individuals for biometric processing.

In comparison, the scheme 3 has not only enrolled a big number of individuals compared to the other two schemes but also achieved a comparable recognition rates.

Experimentally, it can be seen that Scheme 3 has several advantages such as enlarging user population coverage and reducing enrolment failure. It is also observed that small fraction of the population may be unstable for automatic identification.

Table 5.5 Rank-one recognition rates for forehead region using different skin purity assessment schemes (purity threshold = 98%)

Purity scheme	No. of users	LBP			Gabor wavelets		
		1NN	SRC	SVM	1NN	SRC	SVM
Scheme 1	155	0.68	0.73	0.85	0.24	0.45	0.62
Scheme 2	104	0.69	0.83	0.86	0.26	0.36	0.66
Scheme 3	167	0.65	0.70	0.83	0.24	0.48	0.61

Table 5.6 Rank-one recognition rates for right cheek region using different skin purity assessment schemes (purity threshold = 98%)

Purity scheme	No. of users	LBP			Gabor wavelets		
		1NN	SRC	SVM	1NN	SRC	SVM
Scheme 1	203	0.41	0.39	0.58	0.23	0.40	0.52
Scheme 2	158	0.46	0.47	0.61	0.23	0.33	0.51
Scheme 3	206	0.41	0.38	0.57	0.23	0.40	0.52

While rank-one recognition rate is used to represent the identification performance, several conclusions can be drawn from the results in Tables 5.5, 5.6, 5.7, and 5.8.

Table 5.7 Rank-one recognition rates for left cheek region using different skin purity assessment schemes (purity threshold = 98%)

Purity scheme	No. of users	LBP			Gabor wavelets		
		1NN	SRC	SVM	1NN	SRC	SVM
Scheme 1	198	0.38	0.37	0.52	0.22	0.35	0.47
Scheme 2	145	0.43	0.48	0.58	0.20	0.25	0.49
Scheme 3	201	0.38	0.36	0.52	0.21	0.36	0.46

Table 5.8 Rank-one recognition rates for chin region using different skin purity assessment schemes (purity threshold = 98%)

Purity scheme	No. of users	LBP			Gabor wavelets		
		1NN	SRC	SVM	1NN	SRC	SVM
Scheme 1	203	0.54	0.49	0.79	0.23	0.77	0.71
Scheme 2	222	0.52	0.43	0.78	0.21	0.77	0.71
Scheme 3	229	0.52	0.42	0.78	0.22	0.78	0.71

First, LBP descriptor based on skin features has performed generally better than Gabor wavelets for most facial skin regions. Gabor wavelets features achieved lower performance in some regions such as forehead, and right and left skin regions. This observation suggests that LBP descriptor is more convenient for skin feature extraction.

Second, generally, it can be found that forehead region achieved that highest performance compared to other facial skin regions. A possible reason is that forehead region is less likely to be affected by facial expression. The available information is also able to be captured as this region is bigger than the three facial regions' sizes.

Third, comparing the two cheeks regions, it can be found that cheeks regions are less accurate in biometric identification. Although the performance remains stable for both regions, it is found that, skin information extracted from right cheek achieved higher performance with larger population coverage compared to the left cheek. Based on these

observations, cheek regions are highly affected by heavier noise such as head pose, facial expression, and illuminations.

Forth, it is very interesting that, chin region covers the largest populations among other facial skin regions. Considering the results of both methods of feature extraction (i.e. LBP descriptor and Gabor wavelets), this region achieved a generally more stable performance in comparison to other skin regions. Specifically, Gabor wavelets features achieved more stable performance across all skin purity assessment techniques. This only expectation is that Gabor wavelets work with dark skin-tone or those skin images whose appearance affected by hair roots more than bright skin-tone.

Fifth, given different classifiers, focusing only on LBP skin features, SVM classifier achieved the highest performance compared to the two other classifiers. However, SRC achieved better performances when skin information extracted by Gabor wavelet filters for only chin region.

Finally, Scheme 3 is recommended for skin purity assessment technique. The reason is that besides improving population coverage, the proposed scheme is able to achieve comparable performance to other schemes. This finding tends to support the introduction of skin purity assessment in the proposed skin-based biometric system. However, there is clearly a trade-off between the successes of the biometric stage and the failure-to-enrol rate as a result of skin purity assessment filtering.

Considering the best achievement of identification scenario, and using only Scheme 3, Figure 5.13 shows the CMC curves (rank-n recognition rates) for skin feature extraction methods (LBP-based features and Gabor-based features) with the SVM classifier.

There are two observations from Figure 5.13. First, for Figure 5.13 (a), it can be seen from the CMC curves that using LBP-based skin features, the performances of forehead region outperformed the other facial skin regions. However, Chin region achieved a satisfactory performance at rank-five or above compared to forehead performances. The performances of both cheeks are relatively lower which indicates that the identification performances based on skin features extracted from the two cheek regions are closely-related to the noise level of captures. Both suffer from the influence of noise and are

also highly affected by head pose, expressions and illumination leading to degraded performances.

Second, from Figure 5.12 (b), it can be seen that, the chin region achieved the highest performance by using Gabor-based features in comparison to other skin regions. The possible reason is that since the chin region is darker and its skin affected by hair roots especially with men, Gabor wavelet method performed better with this type of skin. This observation may suggest that Gabor-based features are preferable for darker skin images.

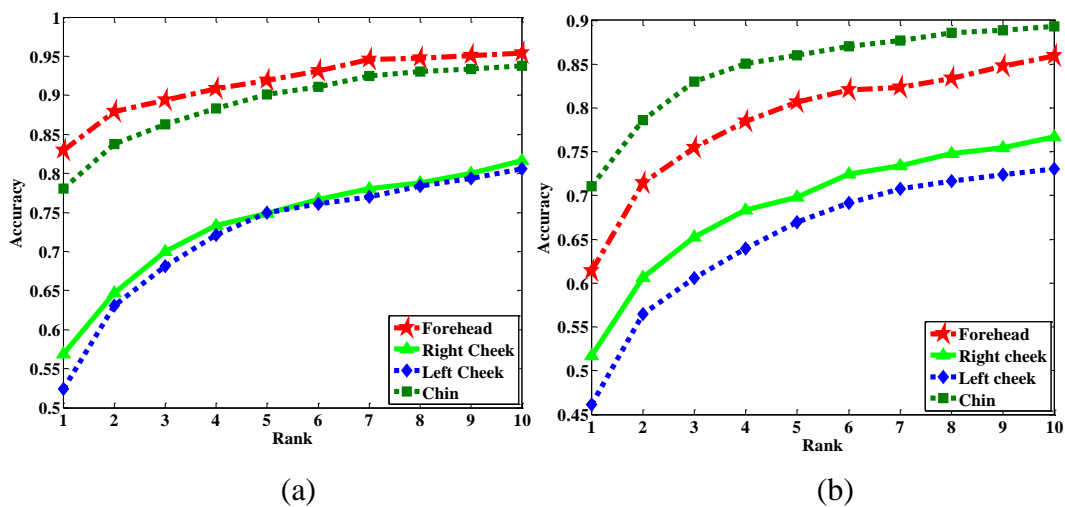


Figure 5.13 CMC curves when using (a) the LBP descriptor and (b) Gabor wavelets with the SVM classifier

5.6.3 Comparative Analysis

For comparison purpose, facial skin regions investigated in this work are extracted from XM2VTS database to be compatible with those in the literature. As Table 5.9 shows the comparative study, the comparison to the competition results is not fully fair. The reason is two-fold. Firstly, the number of individuals used in the evaluation for those in the literature is inaccessible. Secondly, the number of individuals for the proposed system is chosen automatically based skin purity assessment. Thus, it is reasonable to make the comparison because the mechanism used in the present work is closer to the completions.

The analysis in the previous sub-section (Sub-section 5.7.2) is based on the experimental setting in Section 5.6.

Table 5.9 Comparison with other published schemes using the XM2VTS database

Publication	Skin ROI	Skin pixel detection	No. of users	Feature extraction	Classifier	Recognition rate
Lin et al. [5]	Cheeks + Forehead	NA	NA	Regularised LDA	Adaptive metric fusion	0.67
Al-Qarni et al. [184]	Forehead	Manual	84	Gabor wavelets	1NN	0.68
Alsufyani et al.[64]	Forehead	Automatic	98	LBP	1NN SVM	0.76 0.91
<i>Proposed technique</i>	<i>Forehead</i>	<i>Automatic</i>	<i>167</i>	<i>LBP</i>	<i>INN</i> <i>SRC</i> <i>SVM</i>	<i>0.65</i> <i>0.70</i> <i>0.83</i>

However, in this comparison, only the skin biometric performances achieved when Scheme 3 was used for skin purity assessment are compared with the state-of-the-art counterparts. One region is chosen for comparison (forehead region).

As LBP operator is more robust to extract distinct features from the ROI, SVM achieved better performances with this kind of features. According to these observations, the improvement mechanism of proposed skin-based biometric system is compared to existing methods using the same database for evaluation. Table 5.9 shows the rank-one recognition rates for different techniques when using the XM2VTS database.

As compared to other facial skin biometric recognition systems reported in the literature, the improvement of the proposed system covered a larger number of individuals than others with superior rates of recognition accuracy. The only higher recognition rate (0.91) obtained in previous work [64] was only based on 98 individuals; as the proposed system's performance after improving the skin purity assessment scheme 3 was based on 167 individuals, the proposed system correctly identified more people (+49) from the XMT2VTS database despite its slightly lower accuracy.

5.7 Summary

This chapter has proposed a technique for the use in skin-based biometric information for identification. The chapter covered two main set of evaluations.

First, using only the proposed skin pixel detection technique developed in this research, termed Scheme 2 as outlined in (Section 4.5), a novel automated technique is proposed for the biometric recognition of individuals, based on features extracted from the skin. The threshold setting of skin purity assessment was adjusted based on the targeted ROI.

Second, skin purity assessment technique was developed to mitigate the limitation of Scheme 2. Such developed technique (termed Scheme 3) was exploited to improve population coverage besides improving the stability of the skin-based biometric performance. The results of skin purity analysis demonstrate that since Scheme 2

performed better in detecting noisy skin regions, Scheme 3 is preferable for biometric process because of the population coverage.

The proposed automated skin-based biometric system was evaluated to investigate four regions of the facial image and to also establish the most promising region for extraction of identity information. These four regions (forehead, left and right cheek, and chin) were demarcated on the basis of automatically detecting facial landmarks. LBP descriptor and Gabor wavelet method were used to extract skin features, and different classifiers were utilised for classification. Such a technique would be very useful in covert and overt surveillance applications, where full facial images may not be available. Experiments were conducted on the XMT2VTS database with varying noise level. The results suggest that skin texture can provide a rich source of biometric information. The forehead and chin regions were found to provide the most accurate information, probably because these regions are less affected by changes in expression and illumination, and accessories such as spectacles. Comparisons with state-of-the-art algorithms demonstrate the superior performance of the proposed system. Besides enlarging user population coverage and reducing enrolment failure, the proposed system has a competitive, more stable, and maintaining a higher identification performance.

Chapter 6

Usability of Skin Texture Biometrics for Low-Resolution Images

Given the growing demand for biometric technologies that can recognise individuals from skin images at different resolutions (e.g. camera at different distance), this chapter explores the usability of low-resolution facial skin images for biometric purposes. The four facial skin regions investigated here (forehead, right cheek, left cheek, and chin) were selected because at least one of these can usually be captured in real-world scenarios even when part of the face may be occluded. The proposed framework automatically localises and assesses the validity of the region of interest (ROI) for skin-based biometric recognition. Skin features were then extracted using the LBP descriptor because of its effectiveness in extracting skin information. The experiments have been conducted using XM2VTS database. Here, all skin images were resized at different scales to evaluate their utility. The results suggest that skin texture can provide adequate biometric information even when the resolution varies drastically.

For the sake of clarity, it is helpful to distinguish three terms in relation to skin image processing in this chapter: resolution, scale, and distance. The **resolution** of a digital image is generally expressed as the total number of pixels in the horizontal and vertical dimensions. **Image scaling** refers to changes of image size using geometric

transformations. When an image is scaled, the number of pixels changes. This technique is often used to simulate the distance between the captured image and the imaging device. **Distance** affects image resolution; the closer object to the camera, the greater its resolution.

The chapter is organised as follows. Section 6.1 summarises the motivation and the importance of skin image resolution. The experimental setup is described in Section 6.2, and Section 6.3 reports the results. The chapter concludes with a summary in Section 6.4.

6.1 Motivation

This chapter details the use and effectiveness of skin resolution in skin biometric systems. In forensic applications, there is increasing demand for robust face recognition systems ranging from partial to whole face verification. As a standalone source of skin biometric information, a high-resolution face image can be acquired at close range. However, some issues may degrade skin image resolution. For example, if the face is not close enough to the camera or the sensor is of low quality, the image will be small in size, resulting in a low-resolution image that makes recognition difficult. Use of such low-resolution images may cause the system performance be significantly degraded.

Matching two partial facial images of the same subject captured from different distances/resolutions may also be challenging. Because the resolution drop decreases the amount of information available for identifying or verifying individuals [185]. While template images may be high-resolution, a query image is likely to be of low resolution, making matching difficult. With growing numbers of smartphone users and surveillance cameras, partial views of a face can be easily captured at different resolutions. As only a small area of skin may be available, it is therefore very important to be able to exploit this for identification purposes. As explained in Section 3.3, skin texture is the surface texture pattern of any exposed human body part (e.g. face, hand, palm). It has already been demonstrated that skin texture patterns can be utilised for person recognition, especially when using detailed high-resolution images to identify individuals. This chapter investigates both the effect of low-resolution skin images and the possibility of matching these skin images when captured at multiple resolutions.

6.2 Experimental Setup

As partial face recognition at different resolutions has not been adequately addressed to date, the work presented in this chapter investigates the impact of low-resolution images of facial skin texture and scale invariance, and whether such images can be successfully used for biometric recognition/applications. To automatically extract useable skin-based biometric information, the following steps were applied to the ROI: (i) pre-processing, including image normalisation, facial landmark detection and ROI localisation; (ii) skin purity assessment; (iii) feature extraction, and (iv) classification. Each of these steps is elaborated in more detail below and elsewhere in the thesis. For example, the pre-processing stage includes landmark detection (described in Sub-section 5.2.1), region of interest localisation (described in Sub-section 5.2.2), and skin purity assessment technique (described in Sub-section 5.2.3).

LBP histogram has been used for feature extraction as it provided more reliable skin features. In order to match LBP feature vectors provided by multiple operators with different (P, R) values and different skin image resolutions, all feature sets were normalised using the following formula:

$$f'_i = \frac{f_i}{\sum_{i=1}^n f_i} \quad (6.1)$$

where f_i is the feature component before normalising; f'_i is the normalised feature vector; and n is the dimension of the feature vector.

A k -Nearest-Neighbours (k -NN) classifier (with $k=1$) has been used as a common platform for classification.

6.2.1 Generation of Multi Scale Dataste

To simulate the effect of the distance between subject and camera, all facial images in the XM2VTS database were scaled to different resolutions, using interpolation methods.

Commonly used methods of interpolation include nearest neighbour, bilinear interpolation, spline interpolation [186], and bicubic interpolation [187] [188]. Although nearest neighbour and bilinear interpolation seem computationally simple, serious

blurring remains a problem for these methods. It is not obvious which interpolation method is best in terms of skin scale, but bicubic interpolation method is generally considered to achieve better results [172].

To investigate the impact of low-resolution imaging on skin texture information, all facial skin regions were scaled down using bicubic interpolation technique, which essentially involves estimating the values of a pixel at a new position intermediate to the original pixels.

The bicubic method using 16 pixels in the nearest 4x4 neighbours of the original was applied here for scaling of all facial skin images. For all individuals in XM2VTS database, images were then resized to different scales to generate varying resolution images. Specifically, to assess the efficiency of the proposed technique, two datasets (Dataset B and Dataset C) were created by scaling the original images (Dataset A) by a factor of 0.75 and 0.50 respectively. The 0.50 was chosen as the lowest due to the restriction imposed by the LBP parameters (P , R). It also ensures that the ROI would contain sufficient pixels for the LBP operator in both neighbourhoods (P) and radius (R). Figure 6.1 shows some examples of scaled skin images of the forehead region.

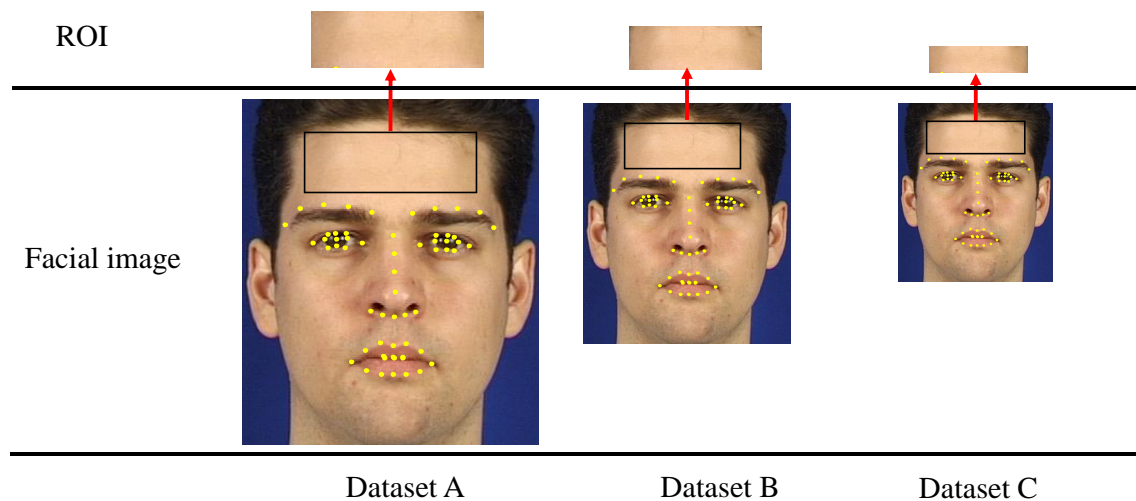


Figure 6.1 Examples of facial skin images and skin regions of interest at different scales

These three datasets were then utilised for biometric processing of each ROI (forehead, right cheek, left cheek, and chin). Table 6.1 shows the size range of the original ROIs and their scales using bicubic interpolation.

The forehead region is larger in size than other skin regions. Both cheek region sizes are smaller, causing feature extraction methods to limit parameters during processing. For example, the cheek region becomes too small (12×12) when resized using scaling factor 0.5.

Table 6.1 ROI size ranges after bicubic interpolation using different scaling factors

ROI	Range size (in pixels)		
	Dataset A	Dataset B	Dataset C
	Original size (from–to)	0.75-scale (from – to)	0.50-scale (from – to)
Forehead	49 x 122–79 x 197	37 x 92–59 x 148	25 x 61–40 x 99
Right Cheek	24 x 24–56 x 57	18 x 18–42 x 43	12 x 12–28 x 29
Left Cheek	24 x 24–56 x 57	18 x 18–42 x 43	12 x 12–28 x 29
Chin	23 x 41–67 x 120	17 x 31–50 x 90	12 x 21–34 x 60

6.3 Experimental Results and Discussion

The rationale of these experiments is based on the fact that facial images are not always of the same resolution when captured at different scales. Variations in resolution may owe to factors such as camera specifications or distance from the camera. For present purposes, ROI was automatically localised in relation to facial landmarks, using Chehra software [20], [177]. In the XM2VTS database, the software could detect landmarks in only 1,128 of the available images, marking all four images for 274 of the 295 available individuals.

The skin purity assessment technique (Scheme 3 as described in Section 4.5) was then applied to each subset to determine the purity of skin images. The number of individuals used here is based on those who passed this scheme (i.e. 167 users as shown in the last row in Table 5.4). Each image that reached the skin purity threshold (see Sections 4.6

and 4.7) was divided into N non-overlapping regions (here, $N = 4$) to extract skin features separately from each of those sub-regions. Finally, all sub-regional features were concatenated to form a single feature vector. Here, skin features were extracted using LBP; as indicated by previous experimental results (see Section 5.7) and also confirmed in many empirical studies, LBP is a powerful and attractive texture descriptor, with excellent results in terms of accuracy and computational complexity [113]. The LBP-based feature vectors were then input to the classifier to be trained, and 1NN classifier was applied to skin texture features.

In the comprehensive set of experiments conducted throughout this thesis to evaluate the performance of the proposed framework, the number of neighbourhoods (P) of LBP parameters was found experimentally to have less influence on biometric performance, but the changes in in number of radius (R) clearly affected performance. Each experiment reported in this chapter explored 8 neighbourhoods ($P = 8$) while differing in radius (R). To match image histograms for different scales, all LBP feature vectors were normalised using Eq. 6.1. The R parameter was adapted to determine the best value of each image scale. The different protocols used for these evaluations are described in the following section.

6.3.1 Experiment 1: The Effect of Resolution

The first experiment had two objectives: (i) to investigate the use of low-resolution skin images to extract biometric information and (ii) to explore the effect of low-resolution images on overall skin-based biometric performance. As outlined in Figure 6.2, in both training and test sets, all images were of the same resolution.

The LBP descriptor extracts skin features from ROI with parameters $P = 8$ and $R = n$, where n varied between 1 and 5. LBP parameters were extracted at different resolutions, and feature vectors were normalised. The experimental protocol used three skin images from each person as a training set; the remaining images were used for testing, with the same R value for both training and test sets.

Table 6.2 shows recognition rates for all facial skin regions at different resolutions, with multiple values of R . For the forehead skin region, 5 was the optimal value of R for the

Dataset A and the Dataset B. The optimal value of R was 3 for the Dataset C indicating that for low-resolution skin images, it is preferable to extract LBP features using a small R value.

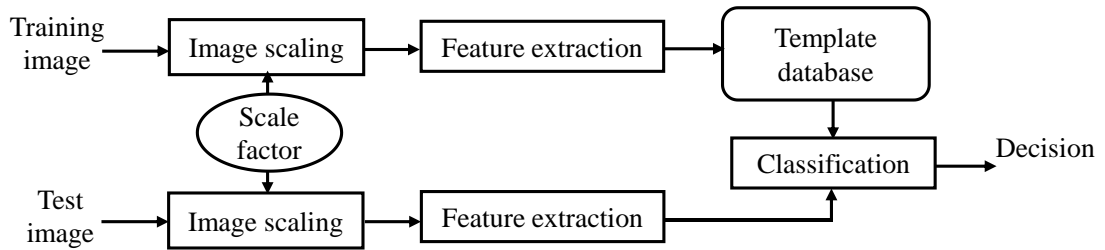


Figure 6.2 Schematic of image scaling framework

Table 6.2 Recognition rates for facial skin regions at different resolutions

ROI	No. of users	Resolution	Recognition rates using LBP(8,R) , 1NN				
			R				
			1	2	3	4	5
Forehead	167	A	0.49	0.53	0.59	0.61	0.63
		B	0.49	0.57	0.63	0.63	0.63
		C	0.50	0.58	0.60	0.55	0.48
Right cheek	206	A	0.33	0.39	0.40	0.39	0.40
		B	0.30	0.38	0.38	0.37	0.29
		C	0.30	0.36	0.27	0.18	-
Left cheek	201	A	0.30	0.35	0.35	0.35	0.35
		B	0.31	0.39	0.36	0.32	0.27
		C	0.30	0.33	0.24	0.14	-
Chin	229	A	0.33	0.40	0.44	0.49	0.48
		B	0.32	0.43	0.45	0.48	0.47
		C	0.33	0.41	0.42	0.40	-

Owing to the size of skin cheek regions, LBP with $R = 2$ delivered the best performance for both. For the cheeks and chin regions, images resized to 0.50-scale (Dataset C) were too small (see Table 6.1), and the R value was set to 2. For the chin region, R was set to 4, as this yielded the best performance for features extracted from Dataset A and Dataset B. These results confirm that LBP parameters play a vital role in extracting discriminant skin information for people recognition, regardless of skin image resolution. Looking more closely, these results indicate that biometric skin information can be obtained using an appropriate feature extraction method even at low resolutions. A planar skin surfaces such as the forehead and chin regions achieved significantly better recognition performance than the less-planar cheek regions.

While recognition rates for the cheek regions were below 0.40, the forehead region achieved a recognition rate of almost 0.60 with different skin image resolutions. This difference of almost 0.20 in recognition rates confirms the utility of the forehead skin region for biometric purposes.

6.3.2 Experiment 2: The Effect of Mismatched Resolution between the Enrolment and Verification

The objective of experiment 2 was to investigate recognition of individuals from skin images at different resolutions. Results for the four facial skin regions are presented at different LBP descriptor parameters.

In real-world scenarios, resolutions of the partial face images captured from different distances or by different devices may often differ from template samples. In this experiment, the training and test sets are examined at different resolutions. Each skin image was assumed to contain distinct skin texture information, as this information was extracted at different scales. To optimise the value of R , different values were analysed.

Figures 6.3 to 6.5 show the recognition accuracy rates achieved using different parameter settings with different evaluation techniques for forehead region. These results suggest that the system performed somewhat better for high-resolution training images (in this experiment, Dataset A or Dataset B) and low-resolution for test images (Dataset C).

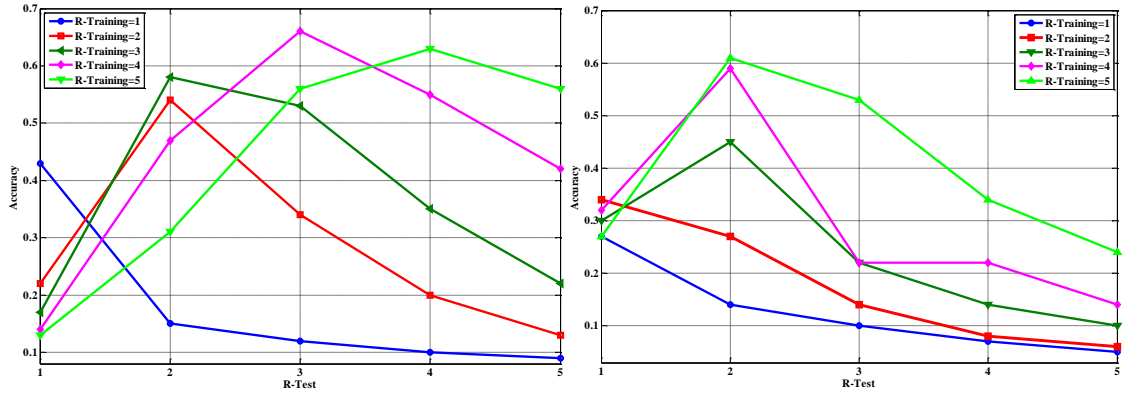


Figure 6.3 Recognition rates with different training and test set R values for matching different resolutions of forehead skin images. The protocol was to train Dataset A and test (a) Dataset B and (b) Dataset C

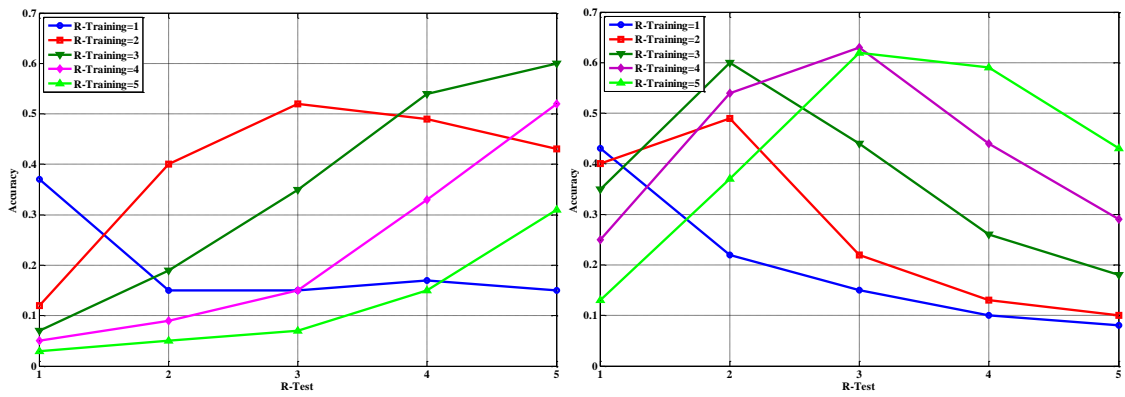


Figure 6.4 Recognition rates with different training and test set R values for matching different resolutions of forehead skin images. The protocol was to train Dataset B and test (a) Dataset A and (b) Dataset C

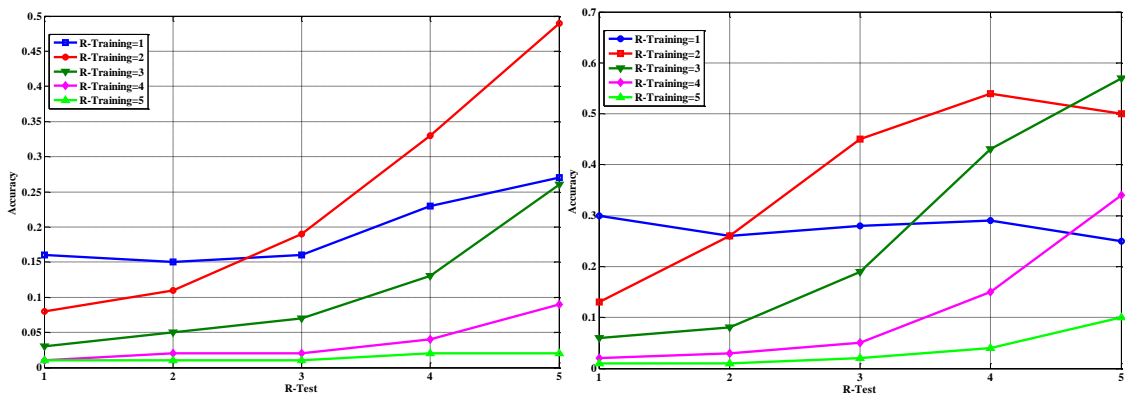


Figure 6.5 Recognition rates with different training and test set R values for matching different resolutions of forehead skin images. The protocol was to train Dataset C and test (a) Dataset A and (b) Dataset B

However, when low-resolution images were used in enrolment and trained (Dataset C), performances deteriorated dramatically. This is possibly because of the lack of discriminatory skin information. In addition, images that were close in resolution (Dataset A vs Dataset B or Dataset B vs Dataset C) yielded better recognition rates. It should also be noted that optimisation of LBP parameters was very important in securing high recognition rates. For example, for the forehead region, high-resolution skin images favour larger values of R (e.g. $R = 4$ or $R = 5$) while low-resolution images favour small values of R (e.g. $R = 2$). These results also suggest that high-resolution images are preferable for enrolment and training for individual recognition.

The three regions (right cheek, left cheek and chin) were also investigated. Each facial skin region of these regions is analysed using a set of experiments of which their results are summarised by extracting the optimal R values. It is clear that the feature extraction is critical in developing a skin-based biometric system for different resolutions. Extracting skin features using only one value of R may not capture sufficient biometric information when the training and test images differ in resolution/distance. In general, it is advisable to extract multiple features to represent skin ROI. For example, to determine the optimal parameters and to capture meaningful features at different resolutions in the training set, it can be argued that, for forehead region, 4 and 5 are the best R values for Dataset A; 3 and 4 are the best for Dataset B; and 2 and 3 are the best for Dataset C. Table 6.3 shows the best R values applied to extract forehead skin features to attain best recognition rates for different resolutions.

Table 6.3 Best recognition rates with different R values for training and test sets when matching images at different scales

ROI	Experimental protocol		R values	Recognition rates	
	Training images	Test images	$(R_{\text{training}}, R_{\text{test}})$		Avg.
Forehead	Dataset A	Dataset B	(4,3)	0.66	0.64
		Dataset C	(5,2)	0.61	
	Dataset B	Dataset A	(3,5)	0.60	0.62
		Dataset C	(4,3)	0.63	
	Dataset C	Dataset A	(2,5)	0.49	0.53
		Dataset B	(3,5)	0.57	

In the preliminary experiments using the forehead skin region to optimise the LBP descriptor parameter, the results suggest that the best approach for the three different resolutions is to extract multiple features (i.e. at least two), using different parameters for each resolution. For example, the Dataset A, $R = 4, 5$; the Dataset B, $R = 3, 4$; the Dataset C, $R = 2, 3$. By training Dataset A, and averaging the accuracies across different test datasets, it achieved higher rank-one recognition accuracies compared to other datasets. Training low-resolution images (Dataset C) achieved a lower identification performance. These parameters were investigated for both cheeks and chin regions and their recognition rates shown in Tables 6.4 - 6.9.

Table 6.4 Recognition rates with different training and test set R values for matching different right cheek resolutions

ROI	Protocol		LBP(8, R) using 1NN classifier		
	Training images	Test images		R_{test}	
			R_{training}	3	4
Right Cheek	Dataset A	Dataset B	4	0.41	0.31
			5	0.29	0.41
		Dataset C	R_{training}	2	3
			4	0.36	0.20
			5	0.35	0.32
		Dataset B	Dataset A	R_{training}	4
	3			0.34	0.37
	4			0.15	0.32
	Dataset C		R_{training}	2	3
			3	0.37	0.25
			4	0.28	0.37
	Dataset C	Dataset A	R_{train}	4	5
			2	0.21	0.30
			3	0.04	0.12
		Dataset B	R_{training}	3	4
			2	0.24	0.30
			3	0.06	0.21

Table 6.4 shows that the accuracy rate for the right cheek region is still around 0.30 for the lowest-resolution images (Dataset C), indicating that most of the skin information

was lost. The best accuracy was achieved when Dataset A was used for training (0.41). The optimal R values are thus listed in Table 6.5. Again, for better recognition, high-resolution skin images are recommended for training. However, despite the varying performance, it can be seen that when Dataset A was trained (original resolution), the performance was more stable compared to other datasets. The rank-one recognition accuracy of right cheek is in the average of 0.41 and 0.37 for testing Dataset B and Dataset C respectively.

Table 6.5 Best recognition rates with different training and test set R values for matching right cheek skin images of different scales (summarised from Table 6.4)

ROI	Experimental protocol		R values	Recognition rate	
	Training images	Test images	($R_{\text{training}}, R_{\text{test}}$)		Avg.
Right Cheek	Dataset A	Dataset B	(5,4)	0.41	0.38
		Dataset C	(4,2)	0.36	
	Dataset B	Dataset A	(3,5)	0.37	0.37
		Dataset C	(3,2)	0.37	
	Dataset C	Dataset A	(2,5)	0.30	0.30
		Dataset B	(2,4)	0.30	

The accuracy of the left cheek region is similar to right cheek indicating that either right or left cheek provided lower accuracy among other facial skin regions. Under such scenarios, the two regions were less applicable for biometric identification, since they were highly affected by pose, light, and expression. The results of evaluation for the left cheek were presented in Table 6.6. The best parameters of LBP were also summarised in Table 6.7.

Table 6.6 Recognition rates with different training and test set R values for matching left cheek skin images of different resolutions

ROI	Protocol		LBP(8, R) using 1NN classifier		
	Training images	Test images	R_{training}	R_{test}	
				3	4
Left Cheek	Dataset A	Dataset B	4	0.40	0.33
			5	0.26	0.36
		Dataset C	R_{training}	2	3
			4	0.37	0.22
			5	0.32	0.31
		Dataset B	Dataset A	R_{training}	4
	3			0.32	0.34
	4			0.14	0.29
	Dataset C		R_{training}	2	3
			3	0.36	0.22
			4	0.26	0.29
	Dataset C	Dataset A	R_{training}	4	5
			2	0.20	0.28
			3	0.04	0.12
		Dataset B	R_{training}	3	4
			2	0.24	0.29
			3	0.06	0.17

Table 6.7 Best recognition rates with different training and test set R values for matching left cheek images of different scales (summarised from Table 6.6)

ROI	Experimental protocol		R values	Recognition rate	
	Training images	Test images	$(R_{\text{training}}, R_{\text{test}})$		Avg.
Left Cheek	Dataset A	Dataset B	(4,3)	0.40	0.39
		Dataset C	(4,2)	0.37	
	Dataset B	Dataset A	(3,5)	0.34	0.35
		Dataset C	(3,2)	0.36	
	Dataset C	Dataset A	(2,5)	0.28	0.28
		Dataset B	(2,4)	0.28	

Table 6.8 Recognition rates with different training and test set R values for matching chin skin images of different resolutions

ROI	Protocol		LBP(8, R) using 1NN classifier		
	Training images	Test images		R_{test}	
			$R_{training}$	3	4
Chin	Dataset A	Dataset B	4	0.47	0.40
			5	0.37	0.48
		Dataset C	$R_{training}$	2	3
			4	0.43	0.27
			5	0.44	0.42
			$R_{training}$	4	5
	Dataset B	Dataset A	3	0.44	0.45
			4	0.22	0.41
			$R_{training}$	2	3
		Dataset C	3	0.44	0.29
			4	0.35	0.45
			$R_{training}$	4	5
	Dataset C	Dataset A	1	0.25	0.21
			2	0.26	0.39
			$R_{training}$	4	5
		Dataset B	2	0.39	0.30
			3	0.30	0.42
			$R_{training}$	4	5

Table 6.9 Best recognition rates with different training and test set R values for matching chin images at different scales (summarised from Table 6.8)

ROI	Experimental protocol		R values	Recognition rate	
	Training images	Test images	$(R_{training}, R_{test})$		Avg.
Chin	Dataset A	Dataset B	(5,4)	0.48	0.46
		Dataset C	(5,2)	0.44	
	Dataset B	Dataset A	(3,5)	0.45	0.45
		Dataset C	(4,3)	0.45	
	Dataset C	Dataset A	(2,5)	0.39	0.41
		Dataset B	(3,5)	0.42	

It is very interesting that chin region has provided better recognition accuracies compared to cheek regions. For the optimal R value, the recognition is almost over 0.40 for all cross dataset resolutions. It can be therefore seen that low-resolution skin images (e.g. Dataset C) is less applicable for training. The best choice for training data is to train Dataset A since it achieved the best recognition rates. Table 6.8 shows the recognition rates with different protocol in terms of training and test data. The summary of the best parameters for this region is presented in Table 6.9.

Finally, it is found that considering all facial skin regions, training multi LBP features is a parameter controlling the trade-off between different resolutions. The features extracted from each LBP parameters mitigate the effect of cross matching resolution error. High resolution images are recommended for training the facial skin features.

6.3.3 Experiment 3: Incorporating Rejections in the Decision Process

In practice, in some cases such as medical diagnosis, or high security organisation, it is often desirable to avoid a wrong decision made by classifiers. If a test sample is wrongly classified, the misclassification could be harmful. Therefore, asking users to present another sample will even cost less than a wrong decision. Classifiers usually assign test samples to a certain class based on its confidence but the classification with a rejection option can abstain the decision due to low classifier confidence. Figure 6.6 shows a common block diagram of rejection technique.

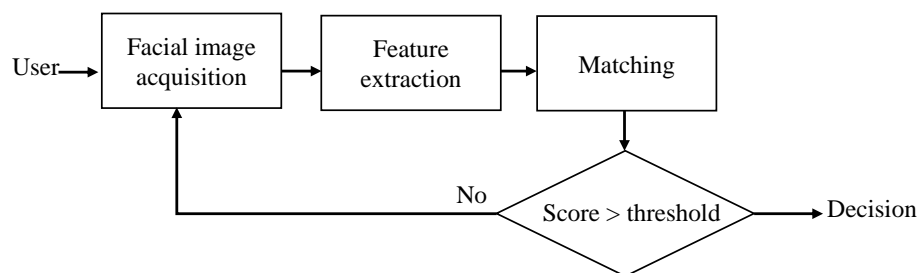


Figure 6.6 Flowchart of error-rejection technique

Rejection option can be incorporated with classifiers to not only improve the robustness of classifiers but also to reduce the error rate. In this experiment, error-rejection curves were analysed for all skin regions investigated in this research. A rejection threshold was applied to the best recognition rates obtained from the best parameters of LBP

descriptor for the four facial skin regions. All these results were presented in Tables 6.4, 6.6, 6.8, and 6.10. Classification error rate is usually defined as the ratio of the number of misclassified test images to the total number of test images. The lower the classification error rate, the better the performance. Error rates of four facial skin regions are shown in Figure 6.7. For better visualisation, in each case, there are six different combinations for evaluation of different skin resolutions. Using the unit interval $[0,1]$ for rejection threshold, error rates are grouped based on facial skin regions where Figure 6.7 (a) shows the error rates for forehead region, Figure 6.7 (b) and (c) show the error rates for both cheek regions and Figure 6.7 (d) shows the error for chin region. As noted above, training set and test set are of different resolutions. The results indicate that using high-resolution skin images for training yielded lower error rates than training with lower-resolution images.

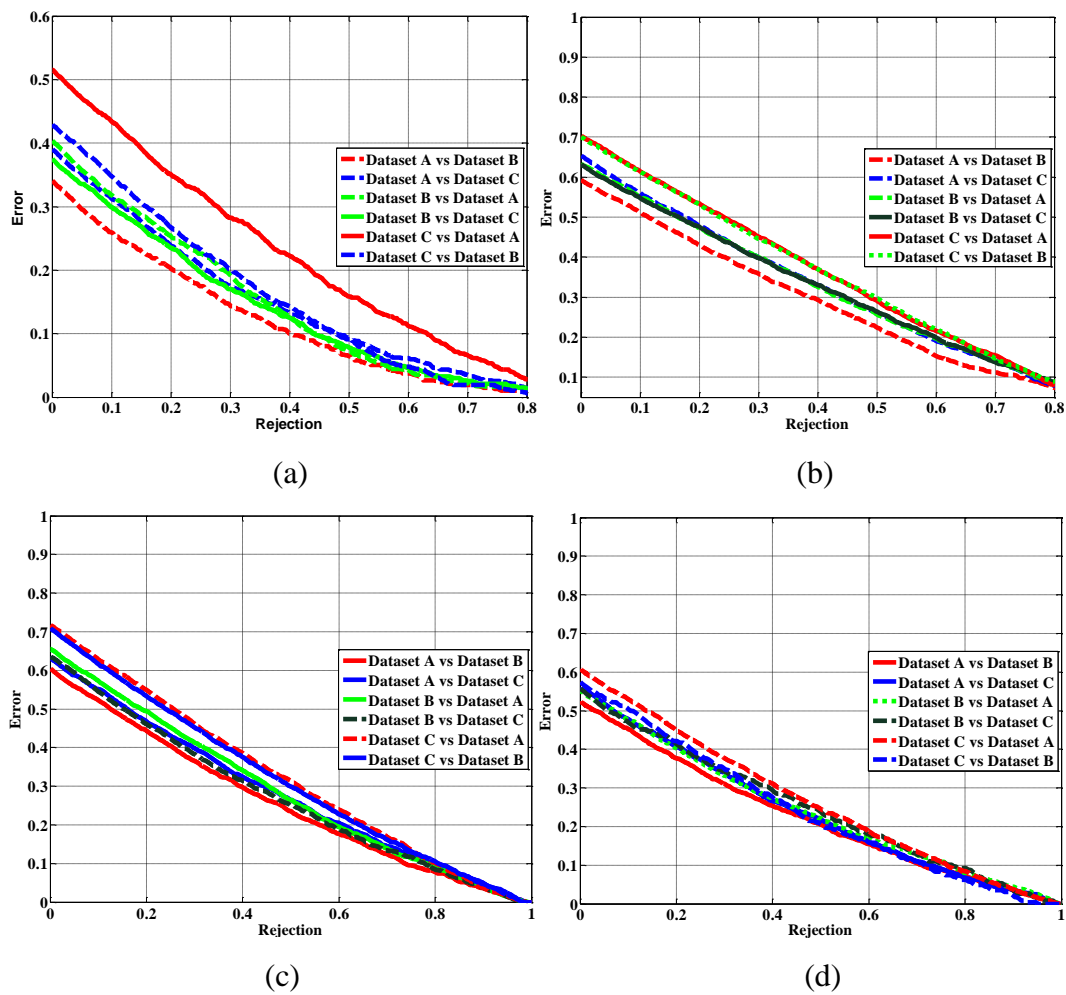


Figure 6.7 Rejection rates for (a) forehead, (b) right cheek, (c) left cheek, and (d) chin region at different resolutions

The performance of high-resolution skin images (Dataset A) is significantly better than the other two datasets on all facial skin regions. It is possible that skin features obtained from high-resolution skin images have more usable and more effective information, leading to less error rates. As low-resolution images (Dataset C) provided higher error rates, more training samples may be needed to train suitable classifiers to improve the performance.

It can be also observed that forehead skin region images have the lowest error rates regardless the resolution of training images. The error rates for both cheek regions are higher than others. Moreover, it is observed that chin region did not only have the large population coverage, but also achieved a promising identification performance.

6.3.4 Experiment 4: Effect of Multi-scale Enrolment using Multiple Parameters

In real scenarios, training and test skin images are unlikely to share the same resolution. This experiment investigated the effectiveness of multi-resolution enrolment with multi-parameters strategy to deal with test images at unknown resolutions. Therefore, multiple templates were created at different scales and different features using multi parameters of LBP-based features. All these set of features are used as a template database during the training phase.

To simulate different skin image resolutions at test phase, the image is down-sampled by a factor of X on XM2VTS database at run time where:

$$0.50 \leq X \leq 1$$

Figure 6.8 shows the flowchart of the proposed technique for this experiment. The trained features stored in the template were constructed by either a specific resolution obtained by any scale factor (i.e. this factor used to create Dataset A, B or C) or by combination of more than scale factor (see Table 6.10 as example of these protocols).

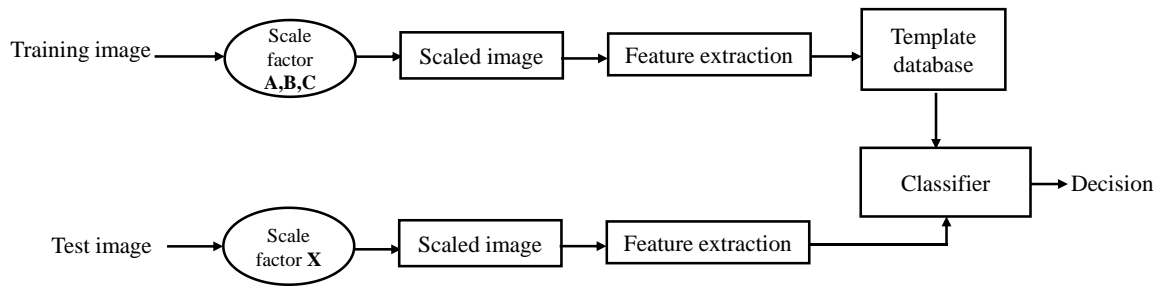


Figure 6.8 Flowchart of the proposed technique for matching different image scales

Different combinations were evaluated to optimise R values for both training and test sets. Tables 6.10, 6.11, 6.12, 6.13 show rank-one recognition rates for different combinations of LBP parameter R in the training and test images used in this protocol. LBP descriptor R values were chosen on the basis of results obtained in the experiment 1. Two main techniques investigated here for training skin features which are: (i) several feature sets extracted at the same resolution but with different R values and (ii) several feature sets are extracted at different skin image resolutions with different R values. Feature sets are then pooled for training so that each test subject may be matched to any of these features as the most similar candidate.

These scenarios were constructed where the training set on the several resolutions (Dataset A, Dataset B, or Dataset C) or a combination of these datasets. The skin features were extracted at multiple LBP parameters R. each of these enrolment scenarios are represented in different columns in Tables 6.10, 6.11, 6.12, and 6.13. For the test images of unknown resolutions, features extracted at different R values are compared with the template database. Each row in Tables 6.10, 6.11, 6.12, and 6.13 shows the identification performances for particular R value(s) used for feature extractions.

Results of identification performances reported in Tables 6.10, 6.11, 6.12, and 6.13 show the effect of different mechanisms for representing skin features. For example, representing two feature vectors obtained from the same skin image resolution or representing multiple feature vectors obtained from different skin image resolutions. The results indicate that capturing multiple image resolution is more robust for the proposed system.

Table 6.10 Forehead region recognition rates (rank-one) for test set images of unknown size based on different combination techniques for training and test sets

Forehead region		Different combinations (+) of training set					
		Several features obtained from the same resolution			Several features obtained from different resolutions		
		Dataset			Dataset		
		A	B	C	A +B	A +B	A +B+C
R values: training set		4,5	3,4	2,3	4,5+3,4	4,5+2,3	4,5 +3,4+2,3
R value(s): test set	2	0.53	0.43	0.53	0.47	0.47	0.47
	3	0.60	0.57	0.42	0.60	0.60	0.60
	4	0.57	0.58	0.52	0.59	0.59	0.60
	5	0.49	0.53	0.52	0.53	0.52	0.55
	2,3	0.53	0.49	0.36	0.54	0.54	0.54
	3,4	0.59	0.57	0.47	0.60	0.60	0.60
	4,5	0.53	0.56	0.53	0.56	0.56	0.57
	3,5	0.55	0.54	0.48	0.57	0.56	0.58
	2,3,4	0.55	0.52	0.42	0.56	0.55	0.56
	3,4,5	0.55	0.56	0.49	0.57	0.57	0.58
2,3,4,5	0.53	0.53	0.44	0.55	0.55	0.56	

Therefore, training multiple skin features extracted from multiple image resolutions is more robust and effective which leads to better trained classifier and hence better performance.

For the test image of unknown resolution, the optimal value of R is 4 because the overall performance for different combination skin resolutions remains more stable compared to other values of R for all facial skin regions. Looking more closely at test images, extracting skin features with a single value of R = 4 or multiple values of R = 3 and R= 4 achieved superior performance.

Comparing the results achieved when the forehead skin features of unknown resolution extracted using $R=4$ and $R=3, 4$, it can be seen from Table 6.10 that training several features obtained from several resolutions achieved the highest performance. This observation indicates that training multi-resolution skin images includes sufficient and more distinctive information than single resolution. The performance is more stable by training combination strategies used for template database (i.e. the recognition rates were around 0.60).

Table 6.11 Right cheek recognition rates (rank-one) for test set images of unknown size based on different combination techniques for training and test set

Right cheek region		Different combinations (+) of training set					
		Several features obtained from the same resolution			Several features obtained from different resolutions		
		Dataset			Dataset		
		A	B	C	A +B	A +C	A+B+C
R values: training set		4,5	3,4	2,3	4,5+3,4	4,5+2,3	4,5 +3,4+2,3
R value(s): test set	2	0.25	0.21	0.13	0.25	0.25	0.25
	3	0.37	0.33	0.23	0.37	0.37	0.37
	4	0.35	0.34	0.28	0.36	0.35	0.37
	2,3	0.30	0.27	0.18	0.31	0.31	0.31
	3,4	0.36	0.34	0.25	0.37	0.36	0.37
	2,3,4	0.32	0.30	0.26	0.33	0.32	0.33

Tables 6.11 and 6.12 show the identification performances for the two cheek regions. The information extracted from these regions is most likely affected by noise such as facial expression, pose, or other artefacts. These factors lead to a drop of performances. Again, the values of $R=4$ and $R=3, 4$ achieved more stable performances compared to other values for unknown test images.

As shown in Table 6.13, chin skin region achieved a better rank-one recognition accuracy compared to the two cheeks. It can also be found that training a multi-resolution image is able to achieve more stable performance. This observation is consistent with the above analysis gained from the other facial skin regions.

Table 6.12 Left cheek recognition rates (rank-one) for test set images of unknown size based on different combination techniques for training and test set

Left cheek region		Different combinations (+) of training set					
		Several features obtained from the same resolution			Several features obtained from different resolutions		
		Dataset			Dataset		
		A	B	C	A +B	A +C	A+B+C
R values: training set		4,5	3,4	2,3	4,5+3,4	4,5+2,3	4,5 +3,4+2,3
R value(s): test set	2	0.23	0.21	0.14	0.23	0.23	0.23
	3	0.33	0.31	0.23	0.34	0.34	0.34
	4	0.33	0.32	0.28	0.34	0.34	0.34
	2,3	0.28	0.26	0.18	0.29	0.28	0.28
	3,4	0.33	0.32	0.25	0.34	0.34	0.34
	2,3,4	0.30	0.32	0.21	0.30	0.30	0.30

Table 6.13 Chin recognition rates (rank-one) for test set images of unknown size based on different combination techniques for training and test set

Chin region		Different combinations (+) of training set					
		Several features obtained from the same resolution			Several features obtained from different resolutions		
		Dataset			Dataset		
		A	B	C	A +B	A +C	A+B+C
R values: training set		4,5	3,4	2,3	4,5+3,4	4,5+2,3	4,5 +3,4+2,3
R value(s): test set	2	0.30	0.25	0.17	0.28	0.30	0.31
	3	0.45	0.40	0.30	0.46	0.45	0.46
	4	0.44	0.44	0.38	0.46	0.46	0.47
	2,3	0.36	0.30	0.23	0.36	0.37	0.38
	3,4	0.44	0.33	0.34	0.45	0.45	0.47
	2,3,4	0.39	0.36	0.28	0.41	0.40	0.41

In conclusion, considering the stability of identification performance, Figure 6.9 shows the CMC curves for the best LBP parameters (i.e. R values) for extracting skin features. Rank-n recognition rates of the combination of the three datasets (A, B and C) with the R = 4 is presented in that figure. Note that rank-one recognition accuracy is shown in the last column in Tables 6.10, 6.11, 6.12, and 6.13.

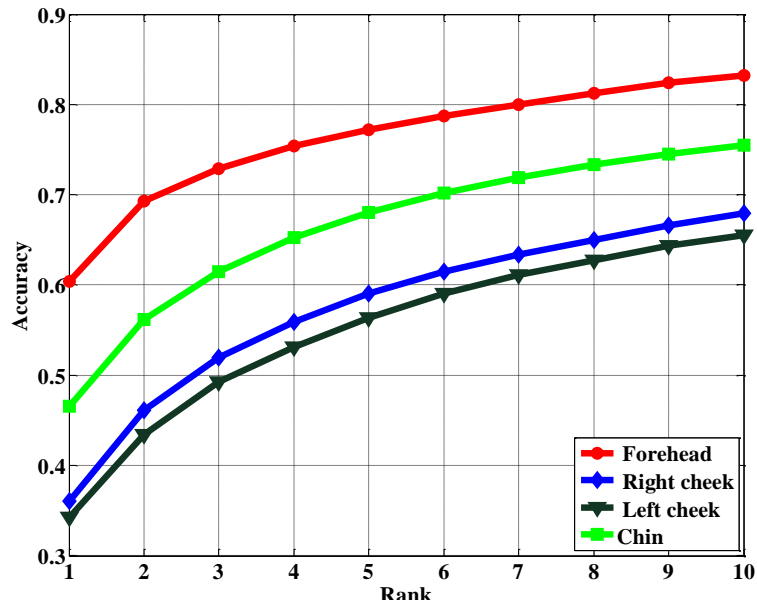


Figure 6.9 CMC curves for the three training set combinations (A, B, and C) with all facial skin regions with $R=4$

6.3.5 Experiment 5: Analysis of Fusion-based Scheme

This experiment explored a multi-classifier fusion approach. Figure 6.10 shows the block diagram for the proposed score fusion technique. Skin features are extracted from the three different skin resolution datasets (A, B, and C) using multiple R values for each resolution. For example, features of Dataset A were extracted with $R = 4, 5$ and feature sets of Dataset B were extracted with $R = 3, 4$. Finally, features of Dataset C were extracted with $R = 2, 3$. Note that these values were chosen based on the experimental results presented above. The outputs of classifiers are fused to a certain way to achieve the final decision. While k -NN is used as a classifier, all scores at match score level were fused to enhance the skin biometric system's performance.

The test image was of random resolution (i.e. scaled by a scale factor X). The chosen R values were $R = 3, 4$ as these values achieved the more stable and the most robust performance in experiment 3. The experiment investigated and applied different score fusion rules in arriving at a final decision. All results are shown in Table 6.14.

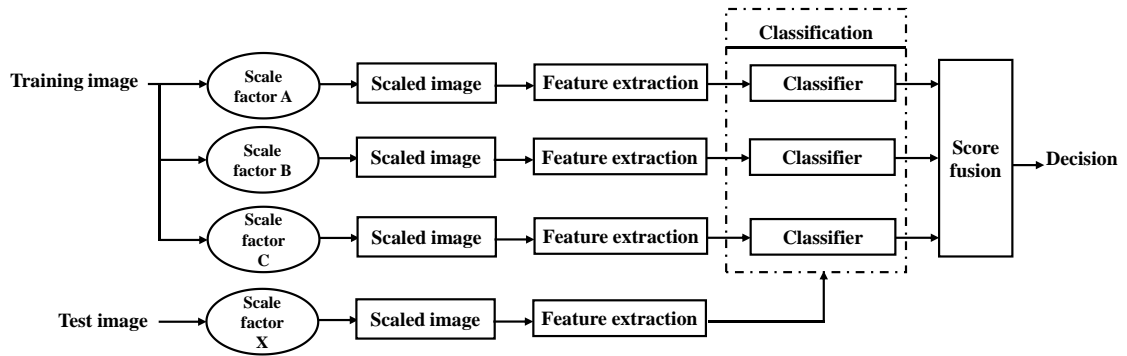


Figure 6.10 Schematic of system framework for score fusion technique

Table 6.14 Recognition rates (rank-one) for test set images of unknown size based on score fusion techniques using k -NN classifier

Score technique	Recognition rates (rank-one)			
	Forehead	Right cheek	Left cheek	Chin
max	0.61	0.37	0.34	0.46
prod	0.59	0.36	0.33	0.46
sum	0.58	0.37	0.34	0.46
min	0.53	0.36	0.33	0.44
median	0.58	0.37	0.34	0.46
vote	0.59	0.37	0.33	0.46

Although all score fusion rules were close to each other and were more stable in identification performance, the max rule delivered slightly better and more stable recognition performance than other fusion rules for all the four facial skin regions. The CMC curve in Figure 6.11 shows rank- n recognition rates for the max rule. From CMC curves, it can be observed that forehead region achieved a superior performances compared to other facial skin regions. Cheek regions had excessively much lower identification performances. It is clear that fusion technique does not always lead to an improved performance.

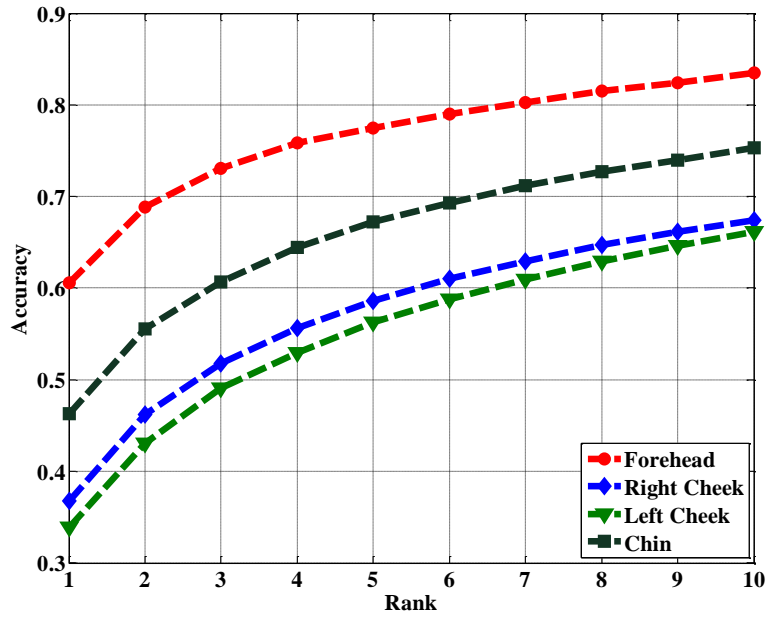


Figure 6.11 CMC curve for the proposed system's recognition rate using max rule

6.3.6 The Performance Comparison between Experiments

To validate the advantages of exploring different techniques for the effect of resolution on skin images for identification performance, a comparison is performed to study the performances between these techniques. Specifically, for experiment 1 and experiment 2, matching different resolutions (experiment 2) does not change much when matching skin images from the same resolution (experiment 1). The other observation is that the training data for experiment 2 is larger than trained data in experiment 1. However, for the experiment 2, the performance was dropped rapidly when Dataset C is trained. Table 6.15 shows the comparative analysis between Experiment 1 and Experiment 2.

Experiment 3 and experiment 4 simulate the scenario where both template database and test images are capture in less constrained environment, and the resolutions of skin images are also different.

Specifically, the technique applied in these experiments simulates the case that skin images in probe can be captured at any resolution. In this case, in skin matching phase, the captured skin image can be different at each time of matching for each individual.

Table 6.15 Results of comparison between Experiment 1 and Experiment 2 for recognition rates

ROI	Experiment 1			Experiment 2			
	Experimental protocol		Recognition rates	Experimental protocol		Recognition rates	
	Training images	Test images		Training images	Test images		Avg.
Forehead	Dataset A	Dataset A	0.63	Dataset A	Dataset B	0.66	0.64
					Dataset C	0.61	
	Dataset B	Dataset B	0.63	Dataset B	Dataset A	0.60	0.62
					Dataset C	0.63	
	Dataset C	Dataset C	0.60	Dataset C	Dataset A	0.49	0.53
					Dataset B	0.57	
Right Cheek	Dataset A	Dataset A	0.40	Dataset A	Dataset B	0.41	0.39
					Dataset C	0.36	
	Dataset B	Dataset B	0.38	Dataset B	Dataset A	0.37	0.37
					Dataset C	0.37	
	Dataset C	Dataset C	0.36	Dataset C	Dataset A	0.30	0.30
					Dataset B	0.30	
Left Cheek	Dataset A	Dataset A	0.35	Dataset A	Dataset B	0.40	0.39
					Dataset C	0.37	
	Dataset B	Dataset B	0.39	Dataset B	Dataset A	0.34	0.35
					Dataset C	0.36	
	Dataset C	Dataset C	0.33	Dataset C	Dataset A	0.28	0.28
					Dataset B	0.28	
Chin	Dataset A	Dataset A	0.49	Dataset A	Dataset B	0.48	0.46
					Dataset C	0.44	
	Dataset B	Dataset B	0.48	Dataset B	Dataset A	0.45	0.45
					Dataset C	0.45	
	Dataset C	Dataset C	0.42	Dataset C	Dataset A	0.39	0.41
					Dataset B	0.42	

As shown in Tables 6.10, 6.11, 6.12, and 6.13 of experiments 3, the best performances were achieved by training different skin resolution images. In experiment 4, the technique is similar but the difference is that score fusion technique was used at score-level fusion to make the final decision. By comparing the results of both experiments,

score-level fusion is less effective in enhancing recognition performance. A possible reason is that in experiment 3 skin images were already trained by extracting features at different resolutions.

6.4 Summary

This chapter investigated the effects of resolution on skin images in different aspects. To that end, all test images were digitally scaled from the original image, using methods such as bicubic interpolation. A set of experiments was evaluated, and results and analyses were reported. The influence of resolution on the skin images was investigated. The results suggest that skin images remain usable as a source of biometric information at lower resolutions. One notable finding was that LBP based on a fixed value of R is not always the best option for multiresolution skin images and scale changes. It can be concluded that multiple skin LBP features are likely to deliver better recognition performance.

Chapter 7

Adaptive Skin Extraction Features Using Biometric Information

7.1 Introduction

The experimental results presented throughout this thesis have demonstrated that the developed skin-based biometric system is capable of providing sufficient biometric information for person recognition. In Chapter 5 and Chapter 6, The ROI was divided into sub-regions to extract skin features from each sub-region. Skin features extracted from these sub-regions were concatenated to form a single feature vector to be classified to identify individuals. This chapter, however, investigates the possibility of classifying skin image using small part of skin region of interest (e.g. identifying individuals from one or more sub-region).

7.2 The Concept and the Motivation

Designing a system for scale-adaptive skin texture features that can exploit the available information from the ROI regardless of the noise this region has is also investigated. In real scenarios, it is difficult to have the same resolution of skin image for both a probe

image and target image. It is also worth to study the use of skin image region of interest at random resolution and at arbitrary location for person identification.

The crucial issue is to extract effective skin texture features from small region of skin image. In this thesis, it has been demonstrated that LBP is an efficient descriptor to describe the local structure of skin texture image with low computational complexity. Therefore, LBP descriptor was only used, in this chapter, for skin features extracted from the facial images of XM2VTS database and classified using several classifiers. The objectives of the experiments presented here were three-fold. First, assessing the performance of the proposed skin biometric system framework with associated issues such as occlusion of the ROI where the individual's recognition based only on a small portion of available skin information (e.g. the ROI may have one sub-region available for feature extraction). Second, adapting the proposed system framework to deal with what available information from the ROI. Third, investigating the characteristics of the skin image being used at random resolutions and arbitrary locations from the ROI for the implementation for person recognition.

Note that, as the forehead yielded the highest accuracy as compared to other facial skin regions, the chapter reports only further experiments on the forehead skin region.

7.3 Experiments

This section presents set of experiments proposed in this chapter for adaptive skin features. Three experiments were mainly conducted on XM2VTS database. The section begins with skin systematic technique when only part of the ROI (only one-sub-region recognition) is used for identification. Then, it describes the method of adaptation of skin features based on extracting available information of ROI. Finally, it introduces a novel technique for recognising individuals using part of facial skin ROI at random resolution from arbitrary locations.

7.3.1 Experiment 1: Skin-based Biometric using Information from each Sub-region

An adaptive skin texture features technique proposed here is based on extracting the LBP features from each sub-region of the entire ROI skin image to identify individuals from that sub-region. At some cases, the ROI might be occluded and only small part of skin image is available. The visible part of skin image could be therefore exploited for person recognition. For this reason, all skin images are partitioned into small equal sizes introducing different number of sub-regions of each skin image. Each sub-region is independently treated (pre-processed) for feature extraction and fed into classifier for identification. In general, the strategy for building the proposed system framework is described as follows:

- i. All ROIs in the database were divided into non-overlapping sub-regions (of equal sizes)
- ii. Each sub-region is passed to skin purity assessment phase (Section 4.5)
- iii. In the training phase, a region is identified as acceptable, it will be submitted to feature extraction stage
- iv. All submitted sub-regions skin images were separately treated to LBP descriptor to extract features from each
- v. Feature vectors extracted from all acceptable sub-regions are accumulated to train the INN classifier
- vi. In the test phase, an acceptable sub-region is submitted to feature extraction stage to be matched with the template database
- vii. For identification scenario, all sub-regions obtained from three images were chosen as a reference gallery set and all sub-regions obtained from the remaining are used as a probe set

Based on the above simulated scenario, the gallery and probe sets are constructed as follows: for each candidate, one sub-region at least was trained of available images. If any candidate has less than one sub-region, they will be excluded to inform that all candidates were enrolled in the template database. Therefore, the number of sub-regions is varied between candidates. For the probe set, the number of sub-images of each candidate was also varied.

The motivation of this experiment is to recognise a person from only one sub-region regardless how many sub-regions are available. The idea is to exploit skin information extracted from small size of the skin image by introducing different numbers of sub-regions of each image. As it has been used throughout this thesis, the skin image was divided into various small non-overlapping sub-regions and each sub-region is independently treated for feature extraction resulting set of feature vectors from the ROI. The number of individuals who passed skin purity assessment has at least one sub-region of its entire available skin image. Table 7.1 shows the number of individuals based on their forehead skin sub-regions which differs of each one and accuracies of the system.

Table 7.1 Recognition rates of single patch obtained from XM2VTS database

Forehead skin region (Total users = 266)			Accuracy
No. of users per available sub-region	No. of an acceptable sub-regions at least	No. of users	LBP with 1NN
10	1	266	-
6	2	256	0.27
3	3	250	0.27
2	4	247	0.28
7	5	245	0.28
2	6	238	0.28
4	7	236	0.28
1	8	232	0.28
7	9	231	0.28
8	10	224	0.29
10	11	216	0.29
13	12	206	0.30
17	13	193	0.31
9	14	176	0.32
33	15	167	0.33
134	16	134	0.37

The first column shows the number of users who have the number of acceptable sub-region(s) presented in the second column at least from the images available in the database. For example, the total number of individuals who have at least one sub-region is equal to 266 and 10 users have only one an acceptable sub-region. Those who have

16 sub-regions (i.e. the four skin ROIs are accepted as pure skin) are only 134. For those who have at least one sub-region, they did not have enough number of sub-regions to be trained as they were excluded from the subsequent analysis. In this case, all individuals were enrolled in the template database. Note that as each person has only four available images in the database, all sub-regions extracted from the three facial images were trained and the remaining sub-regions come from the fourth image were used for testing.

Since the performance of the proposed system is still low, the number of sub-regions being trained has affected the recognition performance. The results also indicate that small sub-regions do not carry much discriminative information. More sub-regions should be trained to improve skin biometric performance. Moreover, it is observed experimentally that the distribution for small size of skin image was overwhelmingly sparse. Figure 7.1 shows the relationship between the number of sub-regions and the recognition rates obtained from this experiment.

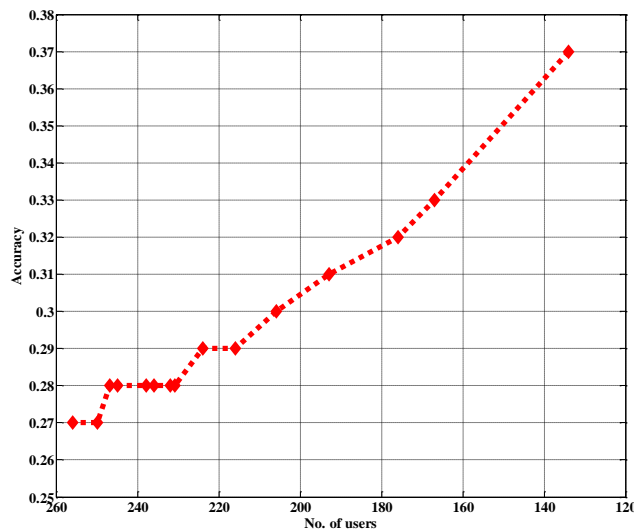


Figure 7.1 Recognition rates for a single sub-region (patch) obtained from forehead skin

7.3.2 Experiment 2: Adaptation of Skin Features based on Extracting Available Information of ROI

The motivation of this experiment is to exploit what available skin information from the region of interest as much as possible.

The experiments carried out through the previous chapters have also rejected any skin region that included a certain noise such as hair. The proposed technique assesses the purity of each skin sub-region from the region of interest (ROI) to be exploited for skin biometric information.

The general systematic diagram of this technique is shown in Figure 7.2. The figure shows the strategy for exploiting the available skin information without excluding the whole skin image if it has a certain amount of noise. The skin region of interest was first divided into m non-overlapping sub-regions with equal sizes. Once each sub-region is assessed, regions, k , passed skin purity assessment scheme were rearranged by concatenating skin sub-regions images to form a new skin region of interest. The new region of interest, I , is then partitioned into m new sub-regions. In this case, as each person has different samples in the database, each image for the same person may differ from the other due to the fact that the reconstructed ROI, I , is based on available sub-regions' locations within original ROI. Thus, it is expected that the texture of the reconstructed region is slightly changed from the original region.

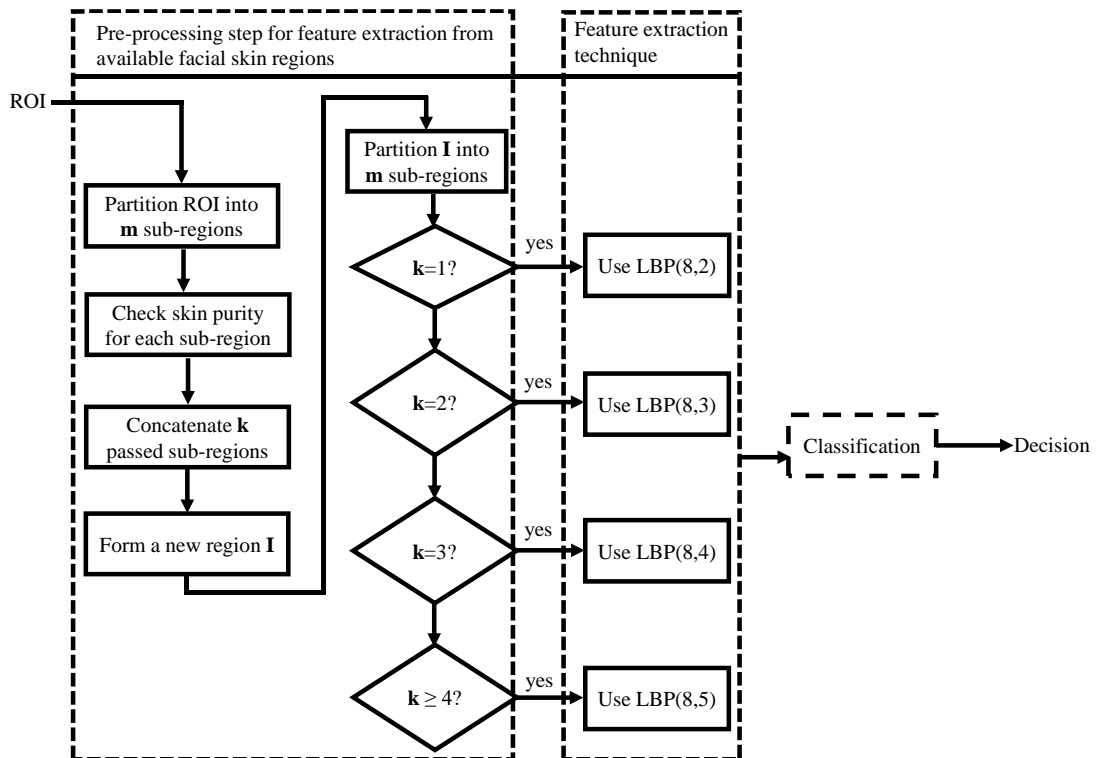


Figure 7.2 Schematic diagram for adaptive skin biometric system

To extract features, the optimal parameters of the LBP descriptor is based on how many sub-regions, k , were classified as an acceptable using skin purity assessment scheme. As presented in Figure 7.2, the parameters (P , R) of LBP descriptor are determined based on the number of pure skin sub-regions passed the skin purity assessment stage.

The experiment was carried out using XM2VTS database to evaluate the proposed system. As demonstrated before, the reconstructed region of interest is divided into four sub-regions irrespective of the number of sub-regions passed the skin purity assessment phase. The LBP was applied with different parameter settings (see Figure 7.2). Depending on the number of sub-regions that had been concatenated, if the skin region of interest only has one sub-region, the radius is set to $R = 2$. The $R = 3$ with two sub-regions, $R = 4$ with three sub-regions, and $R=5$ with greater than or equal four skin sub-regions. In this case, it is important in the present context to note that the training and the test set may vary in the sense that skin features may be extracted by different parameters using LBP descriptor. To put it simply, while features of one subject in the training set may be extracted from one sub-region with one value of R , the test set may be extracted from two or three sub-regions with different values of R (and vice versa).

The recognition rates were presented in Table 7.2 for the reconstructed forehead region of interest. The number of individuals who have at least one sub-region available from each sample in the database (XM2VTS) is 238.

Table 7.2 Recognition rates (rank-one) for forehead region based on LBP descriptor using 1NN and SVM classifiers

No. of sub-regions at least in each image	No. of users	LBP	
		1nn	SVM
		Accuracy	
1	238	0.39	0.53
2	218	0.41	0.58
3	184	0.45	0.62
4	134	0.53	0.70

Considering the LBP performance and the recognition rates are 0.39 and 0.53 using 1NN and SVM classifiers respectively. If the system is more restrict accepting only those who have at least 3 sub-regions, 54 users out of 238 users were rejected and the identification performances are improved to 0.45 and 0.62 for both classifiers 1NN and SVM respectively. SVM provided a significant performance compared to 1NN classifier which is consistent with the results reported in Chapters 5 and 6. As expected, the highest performance was achieved when all sub regions were adequate for biometric process. Considering the performance when applying SVM, it leads to an improvement of 0.17 compared to performance with one-sub-region trained. Figure 7.3 shows CMC curves with rank-n recognition rates being obtained from original LBP using two classifiers, 1NN, and SVM. These results indicate that although texture of skin was reconstructed and its shape is slightly changed due to sub-region movements, skin texture still provides useful biometric information. It can be also seen from Figure 7.3 that the proposed system achieved top performance when the system accepts only those who have four acceptable sub-regions. SVM achieved better performance compared to 1NN.

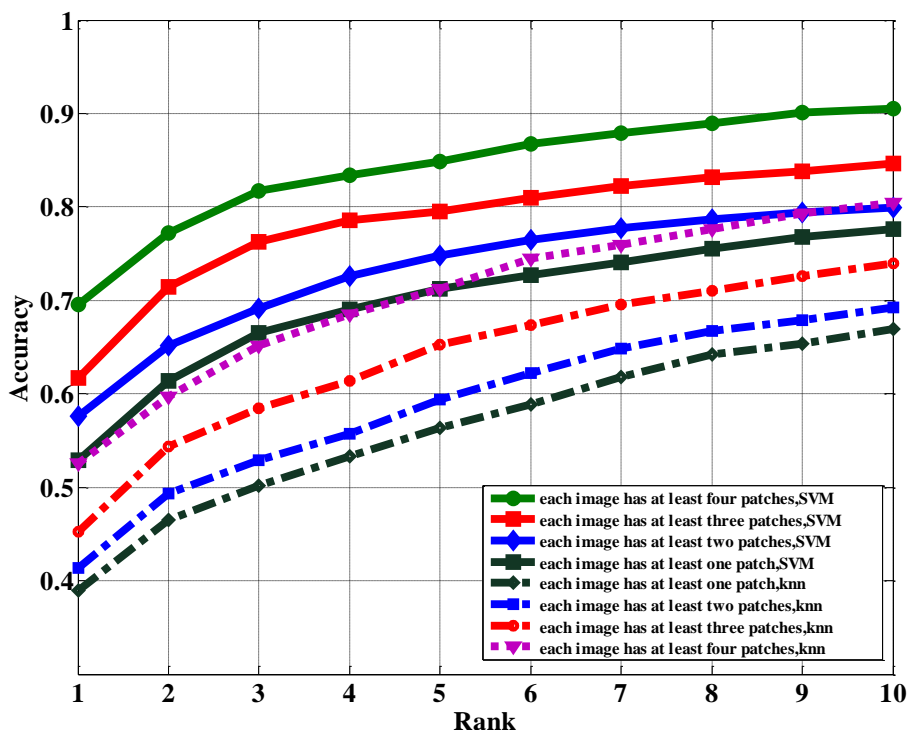


Figure 7.3 CMC curves of rank-n recognition rates of different sub-regions available for feature extraction

7.3.3 Experiment 3: Adaptive Skin Feature Scheme at Arbitrary Scenarios for Skin-based Recognition System

The lack of resolution of imaging systems impacts the performance of biometric systems, especially in the case of surveillance. In real-life applications, the variation in imaging devices may also generate low-resolution image versions of the scene of interest. It is common to search in the gallery and probe for images containing the same person in a cross-camera mode, leading in differing of the resolution for both an enrolment image and a test image. Therefore, the main objective of this experiment is to investigate the ability of matching arbitrary part of skin images of the same ROI in different resolutions in both the template database and the test image.

In order to achieve skin image resolution invariance, multiple features were extracted from each ROI image during the training phase. Note that while the entire ROI is enrolled, the training images can be one of the three different resolutions (Dataset A, Dataset B, Dataset C as explained in Section 6.2) or combination of these datasets. The test images were scaled using an arbitrary scale factor (i.e. the technique was described in Sub-section 6.3.4). Then, part of this image was segmented at arbitrary location of the ROI. The extracted part was used for comparison with those in the gallery set.

The experiment explores the possibility of identifying a person from: (i) unspecified image resolution, and (ii) arbitrary location of the region of interest. The technique was to train single or/and multiple resolution skin images by extracting skin features using multiple LBP parameters. Skin images for training were always selected from single or set of resolution images. The test skin image was, however, randomly scaled and segmented from arbitrary location of ROI for classification. To randomise the test image resolution, the scale factors selected were in the range between 0.5 and 1 of the original skin image size (Dataset A). This ensured occurrence of all possible skin resolutions in the test dataset. A scale factor smaller than 0.5 were not used; otherwise the resulting ROI images would be too small for any viable features extraction using LBP-based method. It should be noted that Dataset C is already of 50% scale showing that when the test image is from this dataset, the effective scale factor can be as low as 0.25 of the Dataset A.

Three different techniques were investigated to train features which were as follows: first, training features extracted from a single dataset (e.g. Dataset A, B, or C). Second, training features extracted from any two datasets and pooled for template database. Third, training features extracted from the three datasets (i.e. Dataset A, Dataset B and Dataset C) and stored in the template database for matching. The justification for applying these settings is that in such scenarios, it is most likely to capture images of persons in less constrained environments. The absence of user's cooperation may also affect the resolution of both training and test image which leads to variation in matching different image resolutions.

Table 7.3 shows the recognition rates for the proposed techniques using different resolutions for training and test images. From this table, it can be seen that four scenarios were evaluated to assess the performance of arbitrary partial view of skin region of interest at random resolution. These scenarios are summarised as follows: the first and the second scenario were to segment randomly at least $\frac{1}{4}$ or $\frac{1}{2}$ of the full size of the ROI from images in Dataset A (termed Scene 1 and Scene 2) respectively. The third scenario was to segment randomly at least $\frac{1}{2}$ of the full size of the ROI from images in Dataset B (termed Scene 3) and the fourth scenario was to segment randomly at least $\frac{1}{2}$ of the full size of the ROI from images in Dataset C (termed Scene 4).

As the smallest scale factor was 0.5 (i.e. used to generate Dataset C), the scale skin images could not be further reduced because the LBP parameters do not accept smaller image sizes. Different R values were explored with each technique and P is set to the most common 8-neighbour form.

The results in Table 7.3 indicate that it is crucial to choose an optimal value of R in such scenarios. It has been observed that in order to deal with different skin resolutions, skin features should be extracted with multiple R values. For example, if the test image is of low resolution, skin features should be extracted using a small R value (and vice versa). As reported in Section 6.3, features extracted from high-resolution images (Dataset A) were recommended to have R values of 4, and 5 while the best R value for the lower resolution (Dataset C) was equal to 2. It is clear that the recognition rates fell when quarter-sized test images were segmented (Scene 1) for matching with Dataset A. In this case, the recognition rates were ranged from 0.22 to 0.26. However, when segmenting

half-sized images (Scene 2) from images in Dataset A, the system performance improved rapidly to just under 0.45. The best recognition rate was achieved for R values of 4 and 5 for both training and test sets. Surprisingly, the proposed system's performance remained constant, with recognition rates of almost 0.40s for the lowest skin resolution (Dataset C) of half-sized images (i.e. Scene 4).

Table 7.3 Recognition rates using part of ROI available for feature extraction

Train		Test image							
		Scene 1		Scene 2		Scene 3		Scene 4	
Training images	R values for training	R values	Accuracy	R values	Accuracy	R values	Accuracy	R values	Accuracy
Dataset A	4	4	0.25	4	0.44	3	0.42	2	0.41
		3,4,5	0.24	5	0.40	2,3	0.41		
	5	4,5	0.23	5	0.41	4	0.43		0.40
		4	0.22	4,5	0.36	3,4	0.41		
	4,5	4,5	0.24	5	0.43	3	0.44		0.41
		4	0.26	4,5	0.44	3,4	0.41		
Dataset B	3	4	0.23	5	0.39	3	0.43		0.40
		5	0.24	4,5	0.39	3,4	0.38		
	4	5	0.18	4	0.26	4	0.40		0.37
		4,5	0.17	5	0.35	3,4	0.33		
	3,4	4	0.24	4	0.40	4	0.41		0.41
		5	0.23	4,3	0.40	3,4	0.40		
Dataset C	2	4	0.16	5	0.31	4	0.36	0.39	
		5	0.18	4,5	0.27	3,4	0.32		
	3	4	0.08	5	0.17	4	0.29	0.38	
		5	0.11	4,5	0.15	3,4	0.21		
	2,3	5	0.17	5	0.28	4	0.34	0.39	
		4,5	0.17	4,5	0.27	3,4	0.34		
Dataset A + B	4,5+3,4	4	0.27	4	0.44	3	0.47	0.43	
		4,5	0.26	4,5	0.42	3,4	0.45		
Dataset A + C	4,5+2,3	4	0.26	4	0.45	3	0.46	0.44	
		4,5	0.26	4,5	0.42	3,4	0.44		
Dataset A + B+ C	4,5+3,4+2,3	4	0.26	4	0.44	3	0.44	0.47	

The best accuracy achieved was 0.47 when the three feature sets obtained from the three datasets (i.e. Dataset A, Dataset B and Dataset C) were trained. To conclude, as expected, extracting features from multiple image resolutions yielded the more stable performance across the four scenes which can be seen clearly in the last row in Table 7.3. This observation is consistent with the demonstrations in previous results in Chapter 6.

7.4 Summary

This chapter investigates the possibility of extracting features from only a small portion of the skin image. The results suggest that discrimination decreases with a smaller amount of information. This chapter also describes a novel method for adaptive skin texture recognition. Using the XM2VTS database, experiments confirmed that the adaptive scheme achieved reasonable performance. This method can be used to exploit any available skin texture information for biometric recognition systems. The other advantage of such system is that enlarging user population coverage and reducing enrolment failure. Skin texture information also achieved reasonable performance when images at random resolutions were arbitrary segmented from the ROI. The proposed system proved capable of recognising individuals even when the skin image was of low resolution.

Chapter 8

Conclusions and Future Work

In this chapter, the conclusions that have been drawn from the research programme are summarised and the recommendations for future work in the field are presented.

8.1 Conclusions

The research work presented in this thesis is concerned with the development of a methodology for the automatic skin-based biometric recognition framework. A prototype of the proposed framework was designed and implemented for skin texture biometric information. Comprehensive experimental analyses were performed for skin-based biometric system framework. The results demonstrate that the proposed system framework have achieved generally improved performance in comparison to state-of-the-art as well as improving population coverage, increasing the degree-of-freedom and reducing the failure-to-enrol rates.

8.2 Summary of Findings

This thesis is concerned with the usability and practicality of facial skin-based biometric technology for wider application in forensic, surveillance, and security contexts where a system must be capable of identifying individuals in crowded scenes. In such scenarios,

the face image may be partly occluded by objects or accessories (sunglasses, hat, scarf, etc.), or by facial hair. As a composition of micro-patterns, skin feature information is a useful source of biometric information for person recognition in such cases. On that basis, the present thesis addresses the design of a robust framework to improve an automatic skin-based biometric system and explores the effect of skin resolution changes on recognition performance.

An improved technique was developed for skin pixel detection. The system was developed to automate the entire process of skin-based biometric system. Several colour-based space techniques were investigated and compared with the state-of-the-art methods. The developed skin pixel detection scheme was evaluated and validated using different databases namely Skin Segmentation dataset and XM2VTS. The experimental results on these benchmark datasets demonstrate that the proposed system shows promise in terms of measures such as Recall and Precision. The technique outperformed all comparison methods in detecting noise with potential distortions in facial skin region. The developed system also promises to remove the need for human interaction and is likely to be more accurate.

Next, a novel automated technique is proposed for skin-based biometric recognition. Four isolated facial skin regions were investigated using the proposed automated system to establish the most promising region for identity information. These four regions (forehead, left cheek, right cheek, and chin) were localised on the basis of automatically detected facial landmarks. As stated above, all skin images used to evaluate the system were assessed automatically for biometric purposes and the system performances using LBP-based and Gabor-based features methods were compared. The influences of optimising the parameters of these two methods are studied and analysed. Several machine learning algorithms for classifying skin texture features were applied including k -nearest neighbours (k -NN), sparse representation classifier (SRC), and support vector machine (SVM) and their results were compared. The comparison with the state-of-the-art algorithms indicates that the proposed skin-based biometric system achieves a generally improved performance.

These results demonstrated that skin texture features can provide a rich source of biometric information. As shown in Chapter 5, the forehead and the chin regions were

found to yield the most accurate biometric information compare to other regions. The reason for providing better performance compared to other skin regions may be because their surfaces are more planer and probably these regions are most likely less affected by changes in expression and illumination.

The previous work in the literature had used only high-resolution skin images. The proposed system was investigated not only using high-resolution skin images, but also using low-resolution skin images. Different techniques were employed including feature level fusion for feature extraction step, and score fusion for classification step. The low-resolution skin results suggested that the training skin images are recommended to be of high-resolution. The training skin image features are also recommended to be extracted using multiple LBP features with different values of radius (R).

Skin images at different resolutions were evaluated and classified with both arbitrary locations and arbitrary resolution of skin region of interest to show the robustness of the proposed system. The skin test samples were selected randomly from the set of available samples. The experimental results suggested that skin image can contribute some biometric information which may be useful in some cases.

The key findings of this thesis which could lead to several novel directions to improve the performance of skin-based biometric systems may be summarised as follows:

- Facial skin pixel detection techniques proposed in Chapter 4 are able to automate skin-based biometric systems. This significantly reduces the need for human interaction and provides the potential for automation of the skin-based biometric process resulting in greater accuracy.
- Skin textural information can be used as a source of biometric information for person recognition. Forehead and chin region are more robust and achieved better performances compared to other facial skin regions. Either the forehead or the chin region is able to provide biometric information for a person in partially occluded facial images. This finding is demonstrated by the experimental results in Chapter 5.

- Low-resolution facial skin images are able to provide biometric information. The performance of low-resolution skin-based biometrics can be improved by extracting multiple features at different resolutions. For example, incorporating more information is able to extract more effective features for skin biometric systems. This is demonstrated by experimental results in Chapter 6.
- Reconstructing available skin part in corrupted and heavier skin region of interest may be possible to identify individuals. This technique and experimental results are presented in Chapter 7.

In conclusion, this thesis suggested a solution of partial view of facial image by using facial skin-based regions. Four facial skin regions were investigated namely Forehead, Right Cheek, Left Cheek and Chin. Each region was investigated independently for person identification. It is likely for some of these regions to be visible even when the entire facial image is not available and can be used on its own for establishing the identity of individuals. Skin texture features provided promising results for person identification. Forehead and chin regions generally achieved better performance compared to cheek regions.

8.3 Recommendations for Future Work

The research work presented in this thesis has indicated the practicability and potential of skin texture features for person recognition. There are, however, a number of areas that could benefit further work.

8.3.1 Skin Detection Improvement

The automatic labelling (segmentation) of the components in skin region of interest presented in Chapter 4 does not reach 100% accuracy and can benefit from further improvement. The technique was built on using colour channels including SVH and

YCbCr colour space. A better performance may be achieved by using different features. A feature set at pixel level could be extracted including LBP, entropy, and Gabor filter components. The features could also be extracted using different window surrounds the pixel. All these features could be added to the skin image values investigated in this thesis. Features could be trained using machine learning technique. The technique should be tested under varying illumination condition using different datasets.

8.3.2 Improvement of Localising Facial Skin Regions

It has been realised that the skin-based biometric systems as discussed in Chapter 5 are based on frontal faces, i.e., all facial landmarks are assumed to be automatically determined using Chehra software. However, this software would fail in the case of non-frontal face captured. This is a limitation of the automatic localisation step which can be avoided by segmenting face boundary from the captured image using skin-based technique.

8.3.3 Improvement of Feature Extraction Technique

The size of the skin region of interest has played a vital role in optimising skin extraction methods. The parameters of LBP descriptor and Gabor wavelet filters were affected by the size of ROI which makes it difficult to generalise these parameters for different sizes. Algorithms may be designed to explicitly model the size of the skin region of interest by dividing the skin images into range of certain sizes, (for instance big, medium, small). This could help to optimise the parameters of feature extraction methods. Features extracted using Gabor wavelet filters may be enhanced by exploiting the whole feature vectors instead of using mean and standard deviation. The high dimensionality may be tackled by using principle component analysis (PCA) [189] or generalized discriminant analysis (GDA) [190].

8.3.4 Improvement of Skin Biometric in Less Constrained Conditions

One of the most promising areas of further research that seems encouraging is that skin-based biometric system could be designed for less constrained environments. It is possible to seek to fuse the features from high-resolution images with large-scale skin features from low-resolution images. As presented in Chapters 5, skin images are of high resolutions and this may be less likely to be captured in real-world scenario. Experiments designed in Chapters 6 and 7 show that there is a possibility of using low-resolution skin images to identify individuals. Evaluating the proposed algorithm to validate its performance could be conducted by using different datasets such as Labeled Face in the Wild (LFW) [191]. Cheek regions could be investigated with skin-texture rotation invariant features. It is obvious that the surface of the cheek regions are less-planer and the recognition error may be caused by viewpoint changes, and large illumination variations.

Skin information at low resolution may be exploited from multiple skin texture feature descriptors to recognise individuals. As shown in Chapters 6 and 7, the recognition rates could be improved by extracting different features from skin images. In order to continue development and advance unconstrained biometric recognition technology, researchers must have access to large amounts of relevant training and testing data with reliable ground truth information. Recently, the IARPA Janus Benchmark – Surveillance (IJB-S) dataset [192], [193] was collected including images at low resolutions which may help researchers to exploit data to improve skin-based biometric system. It is worth noting that IJB databases will be released in the public domain in summer 2018.

The more robust approach that may be considered to achieve better performance is to use deep learning approach. Although it is generally believe that deep learning approaches need sufficient data to be trained, the access to a large amount of data are possible. Very Deep Convolutional Networks for Large-Scale Image Recognition (VGG16) [192] may be applied to validate the robustness of the proposed system framework.

Bibliography

- [1] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Transactions on circuits and systems for video technology*, vol. 14, pp. 4-20, 2004.
- [2] A. K. Jain, K. Nandakumar, and A. Ross, "50 years of biometric research: Accomplishments, challenges, and opportunities," *Pattern Recognition Letters*, vol. 79, pp. 80-105, 2016.
- [3] O. G. Cula, K. J. Dana, F. P. Murphy, and B. K. Rao, "Bidirectional imaging and modeling of skin texture," *IEEE Transactions on Biomedical Engineering*, vol. 51, pp. 2148-2159, 2004.
- [4] O. G. Cula, K. J. Dana, F. P. Murphy, and B. K. Rao, "Skin texture modeling," *International journal of computer vision*, vol. 62, pp. 97-119, 2005.
- [5] D. Lin and X. Tang, "Recognize high resolution faces: From macrocosm to microcosm," in *IEEE Computer Society Conference on, Computer Vision and Pattern Recognition*, 2006, pp. 1355-1362.
- [6] J.-S. Pierrard and T. Vetter, "Skin detail analysis for face recognition," in *IEEE Conference on, Computer Vision and Pattern Recognition, CVPR'07*, 2007, pp. 1-8.
- [7] D. Li, H. Zhou, and K.-M. Lam, "High-resolution face verification using pore-scale facial features," *IEEE transactions on Image Processing*, vol. 24, pp. 2317-2327, 2015.
- [8] L. Hong and A. Jain, "Fingerprint enhancement," in *Automatic Fingerprint Recognition Systems*, ed: Springer, 2004, pp. 127-143.

- [9] A. Kong, D. Zhang, and M. Kamel, "Palmprint identification using feature-level fusion," *Pattern Recognition*, vol. 39, pp. 478-487, 2006.
- [10] R. K. Rowe, "Biometrics based on multispectral skin texture," in *International Conference on Biometrics*, 2007, pp. 1144-1153.
- [11] C. Tang, H. Zhang, and A. W.-K. Kong, "Using multiple models to uncover blood vessel patterns in color images for forensic analysis," *Information Fusion*, vol. 32, pp. 26-39, 2016.
- [12] J. Tang, Z. Li, H. Lai, L. Zhang, and S. Yan, "Personalized age progression with bi-level aging dictionary learning," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, pp. 905-917, 2018.
- [13] C. Ding, C. Xu, and D. Tao, "Multi-task pose-invariant face recognition," *IEEE transactions on Image Processing*, vol. 24, pp. 980-993, 2015.
- [14] L. Chen, M. Zhou, W. Su, M. Wu, J. She, and K. Hirota, "Softmax regression based deep sparse autoencoder network for facial emotion recognition in human-robot interaction," *Information Sciences*, vol. 428, pp. 49-61, 2018.
- [15] C. Chen, A. Dantcheva, and A. Ross, "Impact of facial cosmetics on automatic gender and age estimation algorithms," in *International Conference on, Computer Vision Theory and Applications (VISAPP), 2014* 2014, pp. 182-190.
- [16] B. Klare and A. K. Jain, "Face recognition across time lapse: On learning feature subspaces," in *International Joint Conference on, Biometrics (IJCB)*, 2011, pp. 1-8.
- [17] U. Park and A. K. Jain, "Face matching and retrieval using soft biometrics," *IEEE Transactions on Information Forensics and Security*, vol. 5, pp. 406-415, 2010.

- [18] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," *Computer vision and image understanding*, vol. 61, pp. 38-59, 1995.
- [19] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE transactions on pattern analysis and machine intelligence*, vol. 23, pp. 681-685, 2001.
- [20] A. Asthana, S. Zafeiriou, S. Cheng, and M. Pantic, "Incremental face alignment in the wild," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1859-1866.
- [21] D. Li and K.-M. Lam, "Design and learn distinctive features from pore-scale facial keypoints," *Pattern Recognition*, vol. 48, pp. 732-745, 2015.
- [22] T. Suzuki, "Challenges of image-sensor development," in *IEEE International, Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, 2010, pp. 27-30.
- [23] F. K. S. Chan, X. Li, and A. W.-K. Kong, "A Study of Distinctiveness of Skin Texture for Forensic Applications Through Comparison With Blood Vessels," *IEEE Transactions on Information Forensics and Security*, vol. 12, pp. 1900-1915, 2017.
- [24] T. Igarashi, K. Nishino, and S. K. Nayar, "The appearance of human skin: A survey," *Foundations and Trends in Computer Graphics and Vision*, vol. 3, pp. 1-95, 2007.
- [25] S. Astley, "Classification of Breast Tissue by Texture Analysis," in *BMVC91: Proceedings of the British Machine Vision Conference, organised for the British Machine Vision Association by the Turing Institute 24-26 September 1991 University of Glasgow*, 2012, p. 258.

- [26] G. Guo, L. Wen, and S. Yan, "Face Authentication with makeup changes," *IEEE Transactions on, Circuits and Systems for Video Technology*, vol. 24, pp. 814-825, 2014.
- [27] A. Dantcheva, C. Chen, and A. Ross, "Can facial cosmetics affect the matching accuracy of face recognition systems?," in *IEEE Fifth International Conference on, Biometrics: Theory, Applications and Systems (BTAS)*, 2012, pp. 391-398.
- [28] A. Dantcheva and J. Dugelay, "Female facial aesthetics based on soft biometrics and photo-quality," in *Proceedings of ICME*, 2011.
- [29] C. Chen, A. Dantcheva, and A. Ross, "Automatic facial makeup detection with application in face recognition," in *International Conference on, Biometrics (ICB)*, 2013, pp. 1-8.
- [30] N. Batool and R. Chellappa, "Detection and inpainting of facial wrinkles using texture orientation fields and Markov random field modeling," *IEEE Transactions on, Image processing*, vol. 23, pp. 3773-3788, 2014.
- [31] B. Klare and A. K. Jain, "On a taxonomy of facial features," in *Fourth IEEE International Conference on, Biometrics: Theory Applications and Systems (BTAS)*, 2010, pp. 1-8.
- [32] B. Klare, A. A. Paulino, and A. K. Jain, "Analysis of facial features in identical twins," in *International Joint Conference on, Biometrics (IJCB)*, 2011, pp. 1-8.
- [33] L. Shengcai and A. K. Jain, "Partial face recognition: An alignment free approach," in *International Joint Conference on, Biometrics (IJCB)*, 2011, pp. 1-8.
- [34] Y. Su, S. Shan, X. Chen, and W. Gao, "Hierarchical ensemble of global and local classifiers for face recognition," *IEEE Transactions on, Image processing*, vol. 18, pp. 1885-1896, 2009.

- [35] N. Srinivas, G. Aggarwal, P. J. Flynn, and R. W. V. Bruegge, "Facial marks as biometric signatures to distinguish between identical twins," in *CVPR WORKSHOPS*, 2011, pp. 106-113.
- [36] Y. Ke and R. Sukthankar, "PCA-SIFT: A more distinctive representation for local image descriptors," in *Proceedings of the 2004 IEEE Computer Society Conference on, Computer Vision and Pattern Recognition, CVPR 2004*. 2004, pp. II-506-II-513 Vol. 2.
- [37] A. Savran, N. Alyüz, H. Dibeklioglu, O. Çeliktutan, B. Gökberk, B. Sankur, and L. Akarun, "Bosphorus database for 3D face analysis," in *European Workshop on Biometrics and Identity Management*, 2008, pp. 47-56.
- [38] R. Gross, Matthews, I., Cohn, J. F., Kanade, T., & Baker, S., " Multi-PIE," in *Proceedings of the Eighth IEEE International Conference on Automatic Face and Gesture Recognition*, Amsterdam, Netherlands, Sep. 2008, pp. 1-8.
- [39] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," in *IEEE computer society conference on, Computer vision and pattern recognition, CVPR*, 2005, pp. 947-954.
- [40] A. Nurhudatiana, A. W.-K. Kong, K. Matinpour, S.-Y. Cho, and N. Craft, "Fundamental statistics of relatively permanent pigmented or vascular skin marks for criminal and victim identification," in *International Joint Conference on, Biometrics (IJCB)*, 2011, pp. 1-6.
- [41] C. Tang, A. W. K. Kong, and N. Craft, "Uncovering vein patterns from color skin images for forensic analysis," in *IEEE Conference on, Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 665-672.

- [42] J. Xie, L. Zhang, J. You, D. Zhang, and X. Qu, "A study of hand back skin texture patterns for personal identification and gender classification," *Sensors*, vol. 12, pp. 8691-8709, 2012.
- [43] Y. Jizheng, M. Xia, C. Lijiang, X. Yuli, and A. Compare, "Facial expression recognition considering individual differences in facial structure and texture," *IET, Computer Vision*, vol. 8, pp. 429-440, 2014.
- [44] S. Pamudurthy, E. Guan, K. Mueller, and M. Rafailovich, "Dynamic approach for face recognition using digital image skin correlation," in *AVBPA*, 2005, pp. 1010-1018.
- [45] D. Yi, S. Liao, Z. Lei, J. Sang, and S. Z. Li, "Partial face matching between near infrared and visual images in mbgc portal challenge," in *International Conference on Biometrics*, 2009, pp. 733-742.
- [46] A. K. Jain and U. Park, "Facial marks: Soft biometric for face recognition," in *16th IEEE International Conference on, Image Processing (ICIP)*, 2009, pp. 37-40.
- [47] G. F. Al-Qarni and F. Deravi, "Facial Skin Texture as a Source of Biometric Information," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 5, pp. 57-68, 2012.
- [48] R. Weng, J. Lu, and Y.-P. Tan, "Robust point set matching for partial face recognition," *IEEE transactions on Image Processing*, vol. 25, pp. 1163-1176, 2016.
- [49] C.-W. Tan and A. Kumar, "Human identification from at-a-distance images by simultaneously exploiting iris and periocular features," in *21st International Conference on, Pattern Recognition (ICPR)*, 2012, pp. 553-556.

- [50] J. Neves, F. Narducci, S. Barra, and H. Proença, "Biometric recognition in surveillance scenarios: a survey," *Artificial Intelligence Review*, vol. 46, pp. 515-541, 2016.
- [51] M. Ferrara, A. Franco, D. Maio, and D. Maltoni, "Face image conformance to iso/icao standards in machine readable travel documents," *IEEE Transactions on Information Forensics and Security*, vol. 7, pp. 1204-1213, 2012.
- [52] L. An and B. Bhanu, "Face image super-resolution using 2D CCA," *Signal Processing*, vol. 103, pp. 184-194, 2014.
- [53] Z. Wang, Z. Miao, Q. J. Wu, Y. Wan, and Z. Tang, "Low-resolution face recognition: a review," *The Visual Computer*, vol. 30, pp. 359-386, 2014.
- [54] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, "Face recognition by independent component analysis," *IEEE transactions on Neural Networks*, vol. 13, pp. 1450-1464, 2002.
- [55] J. Ruiz-del-Solar and P. Navarrete, "Eigenspace-based face recognition: a comparative study of different approaches," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 35, pp. 315-325, 2005.
- [56] C.-Y. Yang, C. Ma, and M.-H. Yang, "Single-image super-resolution: A benchmark," in *European Conference on Computer Vision*, 2014, pp. 372-386.
- [57] J. Jiang, C. Chen, J. Ma, Z. Wang, Z. Wang, and R. Hu, "SRLSP: A face image super-resolution algorithm using smooth regression with local structure prior," *IEEE Transactions on Multimedia*, vol. 19, pp. 27-40, 2017.
- [58] G. Dedeoglu, T. Kanade, and J. August, "High-zoom video hallucination by exploiting spatio-temporal regularities," in *Proceedings of the IEEE Computer*

Society Conference on Computer Vision and Pattern Recognition, CVPR., 2004, pp. II-151-II-158 Vol.2.

- [59] Y. M. Lui, D. Bolme, B. A. Draper, J. R. Beveridge, G. Givens, and P. J. Phillips, "A meta-analysis of face recognition covariates," in *IEEE 3rd International Conference on, Biometrics: Theory, Applications, and Systems, BTAS'09*. 2009, pp. 1-8.
- [60] W. W. Zou and P. C. Yuen, "Very low resolution face recognition problem," *IEEE transactions on Image Processing*, vol. 21, pp. 327-340, 2012.
- [61] J. Van Ouwkerk, "Image super-resolution survey," *Image and vision Computing*, vol. 24, pp. 1039-1052, 2006.
- [62] K. Jia and S. Gong, "Generalized face super-resolution," *IEEE transactions on Image Processing*, vol. 17, pp. 873-886, 2008.
- [63] U. Park, H.-C. Choi, A. K. Jain, and S.-W. Lee, "Face tracking and recognition at a distance: A coaxial and concentric PTZ camera system," *IEEE Transactions on Information Forensics and Security*, vol. 8, pp. 1665-1677, 2013.
- [64] H. Alsufyani, S. Hoque, and F. Deravi, "Exploring the Potential of Facial Skin Regions for the Provision of Identity Information," in *Proceeding of the 7th IET International Conference on Imaging for Crime Detection and Prevention (ICDP-16)*, Madrid, Spain, 2016, pp. 1-6.
- [65] H. Alsufyani, S. Hoque, and F. Deravi, "Automated skin region quality assessment for texture-based biometrics," in *Seventh International Conference on, Emerging Security Technologies (EST)*, 2017, pp. 169-174.
- [66] B. Heisele, P. Ho, and T. Poggio, "Face recognition with support vector machines: Global versus component-based approach," in *Proceedings. Eighth IEEE International Conference on, Computer Vision, ICCV*, 2001, pp. 688-694.

- [67] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR, 2001*, pp. I-I.
- [68] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors)," *The annals of statistics*, vol. 28, pp. 337-407, 2000.
- [69] I. Craw, D. Tock, and A. Bennett, "Finding face features," in *European Conference on Computer Vision*, 1992, pp. 92-96.
- [70] E. Hjelmås and B. K. Low, "Face detection: A survey," *Computer vision and image understanding*, vol. 83, pp. 236-274, 2001.
- [71] M.-H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: A survey," *IEEE transactions on pattern analysis and machine intelligence*, vol. 24, pp. 34-58, 2002.
- [72] S. Liao, A. K. Jain, and S. Z. Li, "A fast and accurate unconstrained face detector," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, pp. 211-223, 2016.
- [73] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE transactions on pattern analysis and machine intelligence*, vol. 19, pp. 711-720, 1997.
- [74] R. Gross and V. Brajovic, "An image preprocessing algorithm for illumination invariant face recognition," in *International Conference on Audio-and Video-Based Biometric Person Authentication*, 2003, pp. 10-18.

- [75] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE transactions on Image Processing*, vol. 19, pp. 1635-1650, 2010.
- [76] X. Li and A. W. K. Kong, "A multi-model restoration algorithm for recovering blood vessels in skin images," *Image and vision Computing*, vol. 61, pp. 40-53, 2017.
- [77] N. Q. Huynh, X. Xu, A. W. K. Kong, and S. Subbiah, "A preliminary report on a full-body imaging system for effectively collecting and processing biometric traits of prisoners," in *IEEE Symposium on, Computational Intelligence in Biometrics and Identity Management (CIBIM)*, 2014, pp. 167-174.
- [78] A. Nurhudatiana and A. W.-K. Kong, "On criminal identification in color skin images using skin marks (RPPVSM) and fusion with inferred vein patterns," *IEEE Transactions on Information Forensics and Security*, vol. 10, pp. 916-931, 2015.
- [79] F. K. S. Chan and A. W. K. Kong, "A further study of low resolution androgenic hair patterns as a soft biometric trait," *Image and vision Computing*, 2017.
- [80] Q. Xu, S. Ghosh, X. Xu, Y. Huang, and A. W. K. Kong, "Tattoo detection based on CNN and remarks on the NIST database," in *International Conference on, Biometrics (ICB)*, 2016, pp. 1-7.
- [81] B. Julesz, "Visual pattern discrimination," *IRE transactions on Information Theory*, vol. 8, pp. 84-92, 1962.
- [82] L. Nanni, A. Lumini, and S. Brahmam, "Local binary patterns variants as texture descriptors for medical image analysis," *Artificial intelligence in medicine*, vol. 49, pp. 117-125, 2010.

- [83] C. He, S. Li, Z. Liao, and M. Liao, "Texture classification of PolSAR data based on sparse coding of wavelet polarization textures," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, pp. 4576-4590, 2013.
- [84] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, pp. 2037-2041, 2006.
- [85] C. Ding, J. Choi, D. Tao, and L. S. Davis, "Multi-directional multi-level dual-cross patterns for robust face recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, pp. 518-531, 2016.
- [86] L. Ma, T. Tan, Y. Wang, and D. Zhang, "Personal identification based on iris texture analysis," *IEEE transactions on pattern analysis and machine intelligence*, vol. 25, pp. 1519-1533, 2003.
- [87] M. Pietikäinen, A. Hadid, G. Zhao, and T. Ahonen, *Computer vision using local binary patterns* vol. 40: Springer Science & Business Media, 2011.
- [88] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognition*, vol. 29, pp. 51-59, 1996.
- [89] L. Liu, J. Chen, P. Fieguth, G. Zhao, R. Chellappa, and M. Pietikainen, "A Survey of Recent Advances in Texture Representation," *arXiv preprint arXiv:1801.10324*, 2018.
- [90] A. R. Backes, A. S. Martinez, and O. M. Bruno, "Texture analysis using graphs generated by deterministic partially self-avoiding walks," *Pattern Recognition*, vol. 44, pp. 1684-1689, 2011.
- [91] M. Petrou and P. G. Sevilla, *Image processing: dealing with texture* vol. 10: Wiley Chichester, 2006.

- [92] M. Tuceryan and A. K. Jain, "Texture analysis," *The handbook of pattern recognition and computer vision*, vol. 2, pp. 207-248, 1998.
- [93] A. Materka and M. Strzelecki, "Texture analysis methods—a review," *Technical university of lodz, institute of electronics, COST B11 report, Brussels*, pp. 9-11, 1998.
- [94] A. Rosenfeld and A. C. Kak, *Digital picture processing* vol. 1: Elsevier, 2014.
- [95] H. Tamura, S. Mori, and T. Yamawaki, "Textural features corresponding to visual perception," *IEEE Transactions on, Systems, Man and Cybernetics*, vol. 8, pp. 460-473, 1978.
- [96] G. Castellano, L. Bonilha, L. Li, and F. Cendes, "Texture analysis of medical images," *Clinical radiology*, vol. 59, pp. 1061-1069, 2004.
- [97] W. Richards and A. Polit, "Texture matching," *Kybernetik*, vol. 16, pp. 155-162, 1974.
- [98] D. Charalampidis and T. Kasparis, "Wavelet-based rotational invariant roughness features for texture classification and segmentation," *IEEE Transactions on, Image Processing*, vol. 11, pp. 825-837, 2002.
- [99] T. Pavlidis and Y.-T. Liow, "Integrating region growing and edge detection," *IEEE Transactions on, Pattern Analysis and Machine Intelligence*, vol. 12, pp. 225-233, 1990.
- [100] Y. B. Salem and S. Nasri, "Automatic recognition of woven fabrics based on texture and using SVM," *Signal, image and video processing*, vol. 4, pp. 429-434, 2010.
- [101] L. Lepistö, I. Kunttu, J. Autio, and A. Visa, "Rock image classification using non-homogenous textures and spectral imaging," 2003.

- [102] J. Y. Tou, Y. H. Tay, and P. Y. Lau, "A comparative study for texture classification techniques on wood species recognition problem," in *Fifth International Conference on, Natural Computation, 2009. ICNC'09*. 2009, pp. 8-12.
- [103] A. Latif-Amet, A. Ertüzün, and A. Erçil, "An efficient method for texture defect detection: sub-band domain co-occurrence matrices," *Image and vision Computing*, vol. 18, pp. 543-553, 2000.
- [104] A. Kumar and G. K. Pang, "Defect detection in textured materials using Gabor filters," *IEEE Transactions on, Industry Applications*, vol. 38, pp. 425-440, 2002.
- [105] H. Igehy and L. Pereira, "Image replacement through texture synthesis," *Proceedings. International Conference on, in Image Processing, 1997*, pp. 186-189.
- [106] J. Portilla and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *International Journal of Computer Vision*, vol. 40, pp. 49-70, 2000.
- [107] E. P. Simoncelli and J. Portilla, "Texture characterization via joint statistics of wavelet coefficient magnitudes," in *Proceedings. International Conference on, Image Processing, ICIP 98*.1998, pp. 62-66.
- [108] M. Clerc and S. Mallat, "The texture gradient equation for recovering shape from texture," *IEEE Transactions on, Pattern Analysis and Machine Intelligence*, vol. 24, pp. 536-549, 2002.
- [109] W.-L. Hwang, C.-S. Lu, and P.-C. Chung, "Shape from texture: Estimation of planar surface orientation through the ridge surfaces of continuous wavelet transform," *IEEE transactions on image processing*, vol. 7, pp. 773-780, 1998.

- [110] M. Clerc and S. Mallat, "Shape from texture through deformations," *The Proceedings of the Seventh IEEE International Conference on, in Computer Vision*, 1999, pp. 405-410.
- [111] R. M. Haralick and K. Shanmugam, "Textural features for image classification," *IEEE Transactions on systems, man, and cybernetics*, pp. 610-621, 1973.
- [112] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face recognition with local binary patterns," in *European conference on computer vision*, 2004, pp. 469-481.
- [113] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on, Pattern Analysis and Machine Intelligence*, vol. 24, pp. 971-987, 2002.
- [114] T. Ahonen, J. Matas, C. He, and M. Pietikäinen, "Rotation invariant image description with local binary pattern histogram fourier features," in *Scandinavian Conference on Image Analysis*, 2009, pp. 61-70.
- [115] S. Liao, X. Zhu, Z. Lei, L. Zhang, and S. Z. Li, "Learning multi-scale block local binary patterns for face recognition," in *International Conference on Biometrics*, 2007, pp. 828-837.
- [116] D. Gabor, "Theory of communication. Part 1: The analysis of information," *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, vol. 93, pp. 429-441, 1946.
- [117] G. H. Granlund, "In search of a general picture processing operator," *Computer Graphics and Image Processing*, vol. 8, pp. 155-173, 1978.
- [118] J. G. Daugman, "Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, pp. 1169-1179, 1988.

- [119] S.-H. Cha and S. N. Srihari, "On measuring the distance between histograms," *Pattern Recognition*, vol. 35, pp. 1355-1370, 2002.
- [120] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," *ACM computing surveys (CSUR)*, vol. 31, pp. 264-323, 1999.
- [121] E. Pekalska, P. Paclik, and R. P. Duin, "A generalized kernel approach to dissimilarity-based classification," *Journal of machine learning research*, vol. 2, pp. 175-211, 2001.
- [122] T.-K. Kim and J. Kittler, "Locally linear discriminant analysis for multimodally distributed classes for face recognition with a single model image," *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, pp. 318-327, 2005.
- [123] G. J. Edwards, T. F. Cootes, and C. J. Taylor, "Face recognition using active appearance models," in *European conference on computer vision*, 1998, pp. 581-595.
- [124] J. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "Face recognition using LDA-based algorithms," *IEEE transactions on Neural Networks*, vol. 14, pp. 195-200, 2003.
- [125] H. A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," *IEEE transactions on pattern analysis and machine intelligence*, vol. 20, pp. 23-38, 1998.
- [126] M. J. Er, S. Wu, J. Lu, and H. L. Toh, "Face recognition with radial basis function (RBF) neural networks," *IEEE transactions on Neural Networks*, vol. 13, pp. 697-710, 2002.
- [127] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in *BMVC*, 2015, p. 6.

- [128] S. B. Kotsiantis, I. D. Zaharakis, and P. E. Pintelas, "Machine learning: a review of classification and combining techniques," *Artificial Intelligence Review*, vol. 26, pp. 159-190, 2006.
- [129] T. Denoeux, "A k-nearest neighbor classification rule based on Dempster-Shafer theory," *IEEE Transactions on, Systems, Man and Cybernetics*, vol. 25, pp. 804-813, 1995.
- [130] E. Osuna, R. Freund, and F. Girosit, "Training support vector machines: an application to face detection," in *Proceedings.IEEE computer society conference on, Computer vision and pattern recognition*, 1997, pp. 130-136.
- [131] P. J. Phillips, "Support vector machines applied to face recognition," in *Advances in Neural Information Processing Systems*, 1999, pp. 803-809.
- [132] B. Moghaddam and M.-H. Yang, "Learning gender with support faces," *IEEE transactions on pattern analysis and machine intelligence*, vol. 24, pp. 707-711, 2002.
- [133] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, pp. 210-227, 2009.
- [134] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on, Information Theory*, vol. 13, pp. 21-27, 1967.
- [135] V. Vapnik, "Pattern recognition using generalized portrait method," *Automation and remote control*, vol. 24, pp. 774-780, 1963.
- [136] V. Vapnik, "The nature of statistical learning theory, Springer, New York," 1995.

- [137] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, pp. 273-297, 1995.
- [138] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE transactions on Neural Networks*, vol. 13, pp. 415-425, 2002.
- [139] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the fifth annual workshop on Computational learning theory*, 1992, pp. 144-152.
- [140] C. J. Burges, "A tutorial on support vector machines for pattern recognition," *Data mining and knowledge discovery*, vol. 2, pp. 121-167, 1998.
- [141] A. A. Ross and R. Govindarajan, "Feature level fusion of hand and face biometrics," in *Biometric Technology for Human Identification II*, 2005, pp. 196-205.
- [142] A. Ross, K. Nandakumar, and A. K. Jain, "Introduction to multibiometrics," in *Handbook of biometrics*, ed: Springer, 2008, pp. 271-292.
- [143] Y. Liang, X. Ding, C. Liu, and J.-H. Xue, "Combining multiple biometric traits with an order-preserving score fusion algorithm," *Neurocomputing*, vol. 171, pp. 252-261, 2016.
- [144] L. M. Dinca and G. P. Hancke, "The fall of one, the rise of many: a survey on multi-biometric fusion methods," *IEEE Access*, vol. 5, pp. 6247-6289, 2017.
- [145] J. Kittler, M. Hatef, R. P. Duin, and J. Matas, "On combining classifiers," *IEEE transactions on pattern analysis and machine intelligence*, vol. 20, pp. 226-239, 1998.

- [146] R. Brunelli and D. Falavigna, "Person identification using multiple cues," *IEEE transactions on pattern analysis and machine intelligence*, vol. 17, pp. 955-966, 1995.
- [147] J. Daugman, "Combining multiple biometrics," <http://www.cl.cam.ac.uk/users/jgd1000/combine/combine.html>, 2000.
- [148] L. Lam and S. Suen, "Application of majority voting to pattern recognition: an analysis of its behavior and performance," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 27, pp. 553-568, 1997.
- [149] L. I. Kuncheva, *Combining pattern classifiers: methods and algorithms*: John Wiley & Sons, 2004.
- [150] B. Ulery, A. Hicklin, C. Watson, W. Fellner, and P. Hallinan, "Studies of biometric fusion," *NIST Interagency Report*, vol. 7346, 2006.
- [151] S. Paisitkriangkrai, C. Shen, and A. van den Hengel, "Learning to rank in person re-identification with metric ensembles," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1846-1855.
- [152] X. Hu, S. Peng, J. Yan, and N. Zhang, "Fast face detection based on skin color segmentation using single chrominance Cr," in *7th International Congress on, Image and Signal Processing (CISP)*, 2014, pp. 687-692.
- [153] J. P. Wachs, M. Kölsch, H. Stern, and Y. Edan, "Vision-based hand-gesture applications," *Communications of the ACM*, vol. 54, pp. 60-71, 2011.
- [154] J. M. Palacios, C. Sagüés, E. Montijano, and S. Llorente, "Human-computer interaction based on hand gestures using RGB-D sensors," *Sensors*, vol. 13, pp. 11842-11860, 2013.

- [155] P. Vadakkepat, P. Lim, L. C. De Silva, L. Jing, and L. L. Ling, "Multimodal approach to human-face detection and tracking," *IEEE transactions on industrial electronics*, vol. 55, pp. 1385-1393, 2008.
- [156] W. R. Tan, C. S. Chan, P. Yogarajah, and J. Condell, "A fusion approach for efficient human skin detection," *IEEE Transactions on industrial informatics*, vol. 8, pp. 138-147, 2012.
- [157] K. B. Shaik, P. Ganesan, V. Kalist, B. Sathish, and J. M. M. Jenitha, "Comparative study of skin color detection and segmentation in HSV and YCbCr color space," *Procedia Computer Science*, vol. 57, pp. 41-48, 2015.
- [158] S. L. Phung, A. Bouzerdoun, and D. Chai, "Skin segmentation using color pixel classification: analysis and comparison," *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, pp. 148-154, 2005.
- [159] P. Kakumanu, S. Makrogiannis, and N. Bourbakis, "A survey of skin-color modeling and detection methods," *Pattern Recognition*, vol. 40, pp. 1106-1122, 2007.
- [160] A. Gijssenij, T. Gevers, and J. v. d. Weijer, "Computational Color Constancy: Survey and Experiments," *IEEE transactions on Image Processing*, vol. 20, pp. 2475-2489, 2011.
- [161] V. Vezhnevets, V. Sazonov, and A. Andreeva, "A survey on pixel-based skin color detection techniques," in *Proceedings . Graphicon*, 2003, pp. 85-92.
- [162] M. Kawulok, J. Nalepa, and J. Kawulok, "Skin detection and segmentation in color images," in *Advances in Low-Level Color Image Processing*, ed: Springer, 2014, pp. 329-366.
- [163] K. Sobottka and I. Pitas, "Face localization and facial feature extraction based on shape and color information," in *ICIP (3)*, 1996, pp. 483-486.

- [164] R.-L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, "Face detection in color images," *IEEE transactions on pattern analysis and machine intelligence*, vol. 24, pp. 696-706, 2002.
- [165] I. Zaqout, R. Zainuddin, and S. Baba, "Pixel-based skin color detection technique," *Machine Graphics and Vision*, vol. 14, p. 61, 2005.
- [166] G. Yang, H. Li, L. Zhang, and Y. Cao, "Research on a skin color detection algorithm based on self-adaptive skin color model," in *International Conference on, Communications and Intelligence Information Security (ICCIIS)*, 2010, pp. 266-270.
- [167] N. A. bin Abdul Rahman, K. C. Wei, and J. See, "Rgb-h-cbcr skin colour model for human face detection," *Faculty of Information Technology, Multimedia University*, vol. 4, 2007.
- [168] D. J. Sawicki and W. Miziolek, "Human colour skin detection in CMYK colour space," *IET Image Processing*, vol. 9, pp. 751-757, 2015.
- [169] J. Kovac, P. Peer, and F. Solina, "Human skin color clustering for face detection" in *Proceedings The IEEE Region 8 EUROCON 2003. Computer as a Tool. Ljubljana, Slovenia. Sept., 2003*, pp. 144-148. vol. 2.
- [170] R. B. Bhatt, G. Sharma, A. Dhall, and S. Chaudhury, "Efficient skin region segmentation using low complexity fuzzy decision tree model," in *2009 Annual IEEE India Conference*, 2009, pp. 1-4.
- [171] K. Messer, J. Matas, J. Kittler, J. Luetin, and G. Maitre, "XM2VTSDB: The extended M2VTS database," in *Second international conference on audio and video-based biometric person authentication*, 1999, pp. 965-966.
- [172] S. Battiato, G. Gallo, and F. Stanco, "A locally adaptive zooming algorithm for digital images," *Image and vision Computing*, vol. 20, pp. 805-812, 2002.

- [173] J. Brand and J. S. Mason, "A comparative assessment of three approaches to pixel-level human skin-detection," in *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, 2000, pp. 1056-1059 vol.1.
- [174] F. Ö. Çatak, "Classification with boosting of extreme learning machine over arbitrarily partitioned data," *Soft Computing*, pp. 1-13, 2015.
- [175] R. Min, A. Hadid, and J. L. Dugelay, "Improving the recognition of faces occluded by facial accessories," in *IEEE International Conference on, Automatic Face & Gesture Recognition and Workshops (FG 2011)*, 2011, pp. 442-447.
- [176] X. Wei, C.-T. Li, Z. Lei, D. Yi, and S. Z. Li, "Dynamic image-to-class warping for occluded face recognition," *IEEE Transactions on Information Forensics and Security*, vol. 9, pp. 2035-2050, 2014.
- [177] A. Asthana, S. Zafeiriou, S. Cheng, and M. Pantic. (2014). *Incremental face alignment in the wild*. Available: Online Available : <https://ibug.doc.ic.ac.uk/resources/chehra-tracker-cvpr-2014/> (Access date 2015)
- [178] F. Y. Shih and C.-F. Chuang, "Automatic extraction of head and face boundaries and facial features," *Information Sciences*, vol. 158, pp. 117-130, 2004.
- [179] T. P. Weldon, W. E. Higgins, and D. F. Dunn, "Efficient Gabor filter design for texture segmentation," *Pattern Recognition*, vol. 29, pp. 2005-2015, 1996.
- [180] C.-J. Lee and S.-D. Wang, "Fingerprint feature extraction using Gabor filters," *Electronics Letters*, vol. 35, pp. 288-290, 1999.
- [181] C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition," *IEEE transactions on Image Processing*, vol. 11, pp. 467-476, 2002.

- [182] B. S. Manjunath and W.-Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE transactions on pattern analysis and machine intelligence*, vol. 18, pp. 837-842, 1996.
- [183] K.-K. Sung and T. Poggio, "Example-based learning for view-based human face detection," *IEEE transactions on pattern analysis and machine intelligence*, vol. 20, pp. 39-51, 1998.
- [184] G. F. Al-Qarni and F. Deravi, "Facial Skin Texture as a Source of Biometric Information," *International Journal of Signal Processing, Image Processing & Pattern Recognition*, vol. 5, 2012.
- [185] K. Nguyen, C. Fookes, S. Sridharan, M. Tistarelli, and M. Nixon, "Super-resolution for biometrics: A comprehensive survey," *Pattern Recognition*, vol. 78, pp. 23-42, 2018.
- [186] M. Unser, A. Aldroubi, and M. Eden, "Fast B-spline transforms for continuous image representation and interpolation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 13, pp. 277-285, 1991.
- [187] H. Hou and H. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 26, pp. 508-517, 1978.
- [188] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, pp. 1153-1160, 1981.
- [189] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of cognitive neuroscience*, vol. 3, pp. 71-86, 1991.
- [190] G. Baudat and F. Anouar, "Generalized discriminant analysis using a kernel approach," *Neural computation*, vol. 12, pp. 2385-2404, 2000.

- [191] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," Technical Report 07-49, University of Massachusetts, Amherst2007.
- [192] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [193] C. Whitelam, E. Taborsky, A. Blanton, B. Maze, J. Adams, T. Miller, N. Kalka, A. K. Jain, J. A. Duncan, and K. Allen, "Iarpa janus benchmark-b face dataset," in *CVPR Workshop on Biometrics*, 2017.