3D Face Recognition using Hyperspectral Images

Aman Ghasemzadeh

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Electrical and Electronic Engineering

Eastern Mediterranean University
August 2020
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Ali Hakan Ulusoy
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Doctor of Philosophy in Electrical and Electronic Engineering.

Assoc. Prof. Dr. Rasime Uyguroğlu
Chair, Department of Electrical and Electronic Engineering

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Doctor of Philosophy in Electrical and Electronic Engineering.

Prof. Dr. Hasan Demirel
Supervisor

Examining Committee

1. Prof. Dr. Hasan Demirel
2. Prof. Dr. Bilge Günsel
3. Prof. Dr. Fikret S. Gürgen
4. Prof. Dr. Osman Kükrer
5. Prof. Dr. Hüseyin Özkaramanlı

ABSTRACT

Face is one of the most common biometric modalities which is used for identification. In this context, face recognition has gained an important role in biometric applications based on identification systems during the last few decades. Since there is no physical interaction required during recognition or identification, it's easy to deploy and implement. In face recognition, a face is categorized as known or unknown by comparing a face with all the faces in a database. Due to inherent distinct features, human face analysis is one of the most effective methods of identifying individuals. Nowadays, utilizing hyperspectral images in face recognition is one of the most important research topics in biometrics, since they contain additional significant spectral information compared to 2D images which have only information in spatial dimensions (texture and structure). A hyperspectral image is a data cube containing two spatial and one spectral dimension. Hyperspectral image samples are captured by a hyperspectral camera which operates at multiple narrow bands within the visible spectrum and neighboring near-infrared spectra. Hyperspectral imaging provides new prospects for improving face recognition accuracy since they contain information in both space and spectral axes. Hence significant information for each person regarding the skin based on reflected, absorbed and released electromagnetic energy at different wavelengths can be extracted. Additional spectral information which is not embedded in traditional grey/color facial images provides an opportunity to improve the recognition accuracy. Hyperspectral imaging employs spatial and spectral relationship simultaneously, which improves segmentation and classification in the respective applications. Difficulties encountered in visible light-based face recognition systems, such as the variance in orientation, illumination or expressions can be minimized by

employing hyperspectral imaging. Besides these opportunities, hyperspectral images pose some challenges such as low signal to noise ratios, high dimensionality, and data acquisition needs expensive cameras with multiple sampling in visible and near-infrared spectra. Despite mentioned challenges, hyperspectral images contain more independent and significant information obtained from different sub-bands than 2D images. Hence, hyperspectral images represented in 3D-cubes are by far more capable in classification processes, which is also ideal for spoofing attacks.

In this thesis, we propose novel methods for feature extraction for facial hyperspectral image recognition. The main goal of the thesis is to improve the recognition accuracy of hyperspectral face images. In the first method, three different approaches are proposed employing 3D discrete wavelet transform (3D-DWT) to extract features from the subbands generated by discrete wavelet decomposition. Three approaches include 3D-subband energy (3D-SE), 3D-subband overlapping cube (3D-SOC) and 3D-global energy (3D-GE), which extract different feature vector for each approach containing the energy values calculated from different wavelet sub-bands at different levels of decomposition. Feature vectors generated by three different approaches go through a classifier to complete the face recognition task. In the second proposed method, fusion of spectral information into a single 2D image is achieved by band-specific signal to noise ratio (SNR) based weighting. The fusion method assigns weights based on the calculated band-specific SNR values, weighted sum of the bands generate a single 2D face image. Hence, each pixel along spectral axis is fused to a single pixel resulting a 2D output face image for each 3D hyperspectral face cube. In the third method, in order to fuse spectral bands in hyperspectral face cubes, we apply discrete wavelet transform (DWT) to each pixel along the spectral axis consecutively until the spectral vector for each pixel is decimated to a single pixel transforming the 3D input spectral

face image cube into a 2D output image. 2D output images obtained by the second and third methods are processed using principal component analysis method and face recognition is performed with the help of a classifier.

The experimental results reveal that recognition accuracy of all proposed methods by using standard hyperspectral databases outperform alternative hyperspectral face recognition of the state-of-the-art methods.

Keywords: hyperspectral face image, face recognition, discrete wavelet transform, feature extraction, classification, signal to noise ratio.

ÖZ

Yüz, kimlik tanıma için kullanılan en yaygın biyometrik yöntemlerden biridir. Bu bağlamda, yüz tanıma, son birkaç on yılda kimlik tanıma sistemlerine dayanan biyometrik uygulamalarda önemli bir rol oynamıştır. Yüz tanıma, kimlik tanımlama sırasında fiziksel bir etkileşim gerekmediğinden, dağıtımı ve uygulaması kolaydır. Yüz tanımada, bir yüz, veritabanındaki tüm yüzlerle karşılaştırılarak bilinen veya bilinmeyen olarak kategorize edilir. Doğasında farklı özellikler nedeniyle, insan yüzü analizi bireyleri tanımlamanın en etkili yöntemlerinden biridir. Günümüzde, yüz tanımada hiperspektral görüntülerin kullanılması, sadece mekansal boyutlarda (doku ve yapı) bilgi içeren 2D görüntülere kıyasla ek önemli spektral bilgi içerdiğinden, biyometride en önemli araştırma konularından biridir. Hiperspektral görüntü, iki uzamsal ve bir spektral boyut içeren bir veri küpüdür. Hiperspektral görüntü örnekleri, görünür spektrumda ve komşu kızılötesine yakın spektrumlarda birden çok dar bantta çalışan bir hiperspektral kamera tarafından yakalanır. Hiperspektral görüntüleme, hem boşlukta hem de spektral eksenlerde bilgi içerdiğinden yüz tanıma doğruluğunu iyileştirmek için yeni beklentiler sağlar. Bu nedenle, her insan için farklı dalga boylarında yansıtılan, emilen ve salınan elektromanyetik enerjiye dayanan cilt hakkında önemli bilgiler elde edilebilir. Geleneksel gri / renkli yüz görüntülerine gömülmeyen ek spektral bilgiler, tanıma doğruluğunu artırma fırsatı sunar. Hiperspektral görüntüleme aynı anda mekansal ve spektral ilişkiyi kullanır, bu da ilgili uygulamalarda segmentasyonu ve sınıflandırmayı geliştirir. Yönlendirme, aydınlatma veya ifadelerdeki varyans gibi görünür ışık tabanlı yüz tanıma sistemlerinde karşılaşılan zorluklar, hiperspektral görüntüleme kullanılarak en aza indirilebilir. Bu fırsatların yanı sıra, hiperspektral görüntüler düşük sinyal-gürültü oranları, yüksek

boyutluluk ve veri toplama gibi bazı zorluklar ortaya çıkarır ve görünür ve kızılötesine yakın spektrumlarda çoklu örnekleme ile pahalı kameralar gerektirir. Bahsedilen zorluklara rağmen, hiperspektral görüntüler 2D görüntülerden farklı alt bantlardan elde edilen daha bağımsız ve anlamlı bilgiler içerir. Bu nedenle, 3D küplerde temsil edilen hiperspektral görüntüler, kimlik sahtekarlığı saldırıları için de ideal olan sınıflandırma işlemlerinde çok daha yeteneklidir.

Bu tezde, yüz hiperspektral görüntü tanıma için öznitelik çıkarımı için yeni yöntemler öneriyoruz. Tezin temel amacı, hiperspektral yüz görüntülerinin tanınma doğruluğunu arttırmaktır. İlk yöntemde, ayrık dalgacık ayrışımı ile oluşturulan alt bantlardan öznitelikler elde etmek için 3B ayrık dalgacık dönüşümü (3D-DWT) kullanılarak üç farklı yaklaşım önerilmektedir. Önerilen üç yaklaşım, 3B alt bant enerjisi (3D-SE), 3B alt bant örtüşen küp (3D-SOC) ve 3B küresel enerji (3D-GE) olup, farklı dalgacık alt bantlarından hesaplanan enerji değerlerini içeren her yaklaşım için farklı öznitelik vektörü çıkarılır. Üç farklı öznitelik vector sınıflandırıcıdan geçirilerek yüz tanıma işlemi tamamlanmaktadır. Önerilen ikinci yöntemde, spektral bilginin tek bir 2B görüntüye füzyonu, banda özgü sinyal/gürültü oranı (SNR) tabanlı ağırlıklandırma ile elde edilir. Füzyon yöntemi, hesaplanan banda özgü SNR değerlerine dayalı olarak ağırlıklar atanarak, bantların ağırlıklı toplamı tek bir 2B yüz görüntüsü oluşturulur. Üçüncü yöntemde, hiperspektral yüz küplerindeki spektral bantları birleştirmek için, her piksel için spektral vektör 3B giriş spektralini tek bir piksele dönüştürülene kadar spektral eksen boyunca her piksele ayrı dalgacık dönüşümü (DWT) uygulayarak, yüz küpü bir 2B çıktı görüntüsüne dönüştürülür. İkinci ve üçüncü yöntemlerle elde edilen 2B çıktı görüntüleri Ana bileşenler analizi yöntemi ile dönüştürülerek bir sınıflandırıcı yardımı ile yüz tanıma işlemi gerçekleştirilmektedir.

Deneysel sonuçlar, standart hiperspektral veri tabanları kullanılarak önerilen tüm yöntemlerin doğruluğunun, modern yöntemlerin alternatif hiperspektral yüz tanıma işleminden daha iyi performans gösterdiğini ortaya koymaktadır.

Anahtar Kelimeler: hiperspektral yüz görüntüsü, yüz tanıma, ayrık dalgacık dönüşümü, özellik çıkarımı, sınıflandırma, sinyal-gürültü oranı.

DEDICATION

Dedicated to

My beloved parents, my dear husband and my dear supervisor who always supported me during my Doctoral studies.

ACKNOWLEDGMENT

I would like to thank to my supervisor Prof. Dr. Hasan Demirel for his patience, motivation, knowledge and continuous support in my PhD study.

Special thanks to my dear parents for all support and encouragement throughout my study which spiritually helped me to complete my education. This achievement would not have been possible without them.

I would like to thank to my husband Mustafa, for his patience, endless love, encouragement and supports.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

$\alpha(x, y)$	Average of Pixel Along Spectral
γ	Training Set
δ_{high}	Decomposed Hyperspectral Face Cube Passing Through HPF
δ_{low}	Decomposed Hyperspectral Face Cube Passing Through LPF
ε	Energy Vector
$\varepsilon(x, y)$	Energy of Pixel Along Spectral
ε_G^1	Energy Vector of First Level Wavelet
ε_{GLLL}^2	Energy Vector of Second Level Wavelet
λ	Regulation Parameter
λ_k	Eigenvalue
μ	Mean
ψ	Training Set with Zero Mean
ω_n	Weight
Φ	Weighted Linear Combination of All Subbands
Ω_g	Weighted Linear Combination
AUC	Area Under the Curve
CNN	Convolutional Neural Network
cov	Covariance Matrix
CRC	Collaborative Representation Classifier
DWT	Discrete Wavelet Transform
E_g	Energy of Whole Cube
$E_G^1(Z)$	First Level Energy of Global Cube

$E_{GLLL}^{2}(Z)$	Second Level Energy of Global Cube
FFT	Fast Fourier Transform
FN	False Negative
FP	False Positive
FPR	False Positive Rate
FS	Fourier Spectrum
GE	Subband Global Energy
$g(Z)$	Highpass Filter Kernel
HIS	Hyperspectral Image
HOG	Histogram of Oriented Gradients
HPF	High Pass Filtering
HSFD	Hyperspectral Face Database
HW	Hybrid Wavelet
$h(z)$	Lowpass Filter Kernel
k-NN	*K*-Nearest Neighbor
KW	Kekre Wavelet
LBP	Local Binary Pattern
LDP	Local Derivative Pattern
LPF	Low Pass Filtering
NIR	Near Infrared
OWDWT	Overlapping Cube
PCA	Principal Component Analysis
PDWT	Pixel-Based Shift
P_g	Wavelet Coefficient
P_G^1	First Level Wavelet Coefficient

P_{GLLL}^{2}	Second Level Wavelet Coefficient
ROC	Receiver Operating Characteristic
SE	Subband Energy
SNR	Signal to Noise Ratio
SOC	Subband Overlapping Cube
STD	Standard Deviation
TN	True Negative
TP	True Positive
TPR	True Positive Rate
u_k	Eigenvector
W	Subband Cube
WDWT	Non-Overlapping Cube

Chapter 1

INTRODUCTION

1.1 Introduction

Face recognition has gained an important role in biometric application based identification systems during last few decades. In face recognition task, by comparing a face with a database that contains many faces, system can categorize a face as either known or unknown [1]. Identifying facial features, which is divided in two categories (global and local) is one of the most effective methods to analyze the human face. In global category such as the whole face, while in local category such as regions of eyes, nose, and mouth vary significantly in details across each individual [2]. Face recognition systems in visible light have some challenges such as variance in orientation, illumination or face expression which causes the certain variability in embedded information [3]. As most of faces consist of two eyes, mouth and nose which are in the same location, face expression is a difficult problem in face recognition systems. Face recognition comparing to other biometrics such as iris, palmprint and fingerprint recognition has an advantage of non-contacting and no interacting with the person for identification. One way to overcome mentioned challenges is using depth information of 3D face images. Facial hyperspectral images are a new topic in biometrics as it improves the recognition performance according to the significant information included in spectral dimension.

In this thesis, the focus is on the feature extraction and spectral information fusion for recognition and classification of facial hyperspectral images. In the first proposed method, 3D discrete wavelet transform (3D-DWT) is employed to extract features from the subbands generated by discrete wavelet decomposition in three different approaches. Three approaches include 3D-subband energy (3DSE), 3D-subband overlapping cube (3D-SOC) and 3D-global energy (3D-GE), which extract different feature vector for each approach containing the energy values calculated from different wavelet subbands at different levels of decomposition. Feature vectors generated by three different approaches go through a classifier to complete the face recognition task. In the second proposed method, discrete wavelet transform (DWT) is applied to each pixel along the spectral axis consecutively until the spectral vector for each pixel is decimated to a single pixel transforming the 3D input spectral face image cube into a 2D output image. In the third proposed method, band-specific signal to noise ratio (SNR) based weighting is proposed to achieve a single 2D image by fusion of spectral information. The fusion method assigns weights based on the calculated band-specific SNR values, weighted sum of the bands generate a single 2D face image. Hence, each pixel along spectral axis is fused to a single pixel resulting a 2D output face image for each 3D hyperspectral face cube. 2D output images obtained by the second and third methods are processed using principal component analysis method and face recognition is performed with the help of a classifier. In this manner three standard hyperspectral face databases (HSFD) are used. PolyU-HSFD [4], CMU-HSFD [5] and UWA-HSFD [6].

1.2 Facial Hyperspectral Images

In spectral imaging, information is collected and processed through electromagnetic spectrum. In hyperspectral imaging, the spectrum of each pixel is obtained for the

purpose of detection and recognition of objects [7]. In facial hyperspectral images, samples are captured at multiple narrow bands within the neighborhood of visible light in the electromagnetic spectrum, which reveals significant information that is not evident in traditional grey/color images. A hyperspectral image is a data cube with two spatial and one spectral dimension. Hyperspectral camera operates at multiple narrow bands of the visible spectrum and beyond hence hyperspectral imaging cover a wide range of wavelengths. Hyperspectral imaging measures contiguous spectral bands which include spectral details in the visible, near-infrared or ultraviolet bands, hence it imports more information comparing to traditional grey/color images. As the spectrum is divided into many bands, significant information for each person regarding the skin based on reflected, absorbed and released electromagnetic energy at different wavelengths can be achieved [8]. Having many bands, hyperspectral images contains extra significant information in spectral bands, compared to traditional grey/color facial image data [8-9]. Although, the acquisition of hyperspectral images is more sophisticated which involves more expensive cameras with multiple sampling in near-infrared and ultraviolet spectra, the recent improvements in camera technologies and optics made it more economical and plausible. Despite of relatively higher acquisition cost, the amount of independent information coming from different sub-bands is much more than a standard 2D camera. Hence, hyperspectral images represented in 3D-cubes are by far more capable in classification processes, which is also ideal for spoofing attacks. In order to further improve segmentation, recognition and classification of the images, hyperspectral imaging employ spatial and spectral relationship simultaneously. By employing the application of hyperspectral imaging, difficulties encountered in visible light-based face recognition systems, such as the variance in orientation, illumination or expressions can be minimized [3].

1.3 Problem Definition

There are several difficulties and limitations encountered in face recognition such as variations in skin color under different illumination, different face angles and variance in orientation, clarity of the face image according to the distance of subject from camera and face expression [49]. To overcome some of mentioned difficulties and limitations, spectral information of hyperspectral face images can be employed in face recognition systems. In addition to opportunities, hyperspectral images pose some challenges such as low signal to noise ratios in some spectral bands, high dimensionality and data acquisition needs expensive cameras with multiple sampling in visible and near-infrared spectra. Regarding high dimensionality of hyperspectral face images, in almost all existing hyperspectral face recognition techniques there are algorithms which are adopted for dimensionality reduction in the form of feature extraction.

1.4 Thesis Objectives

This thesis work is about overcoming the challenges in face recognition by using spectral information of hyperspectral face images for feature extraction and classification. The main objectives of this research work are listed below:

1. Using 3D Discrete Wavelet Transform (DWT) to isolate 3D data into frequency sub-bands. 3D-DWT decompose volumetric data in horizontal, vertical and depth directions. Hence it is an appropriate decomposition procedure for feature extraction in 3D data.
2. Applying Discrete Wavelet Transform (DWT) for dimensionality reduction and feature extraction. DWT analyze signals in time and frequency domains simultaneously. Hence it is an appropriate procedure to analyze spectral information in facial hyperspectral images

3. Using band-specific signal to noise ratio (SNR) based weighting to fuse spectral information acquired through different spectral bands into a single 2D image.

1.5 Thesis Contributions

In this research work, several methods are proposed to improve the recognition performance. Major contributions of this thesis are listed below:

1. Extracting the wavelet coefficients of whole hyperspectral image cube by 3D-DWT as a feature vector. The texture property is calculated by energy value from different wavelet sub-bands at different levels of decomposition.
2. To avoid edge-blurring effects and spatial information loss, each hyperspectral face image is divided to adjacent cubes and an overlapping cube. To fuse spatio-spectral information in overlapping cubes in each hyperspectral face image, 3D wavelet features are extracted globally.
3. In the *n*-level 3D-DWT at each level, a filtered signal is represented by detail coefficients, which only span half of the frequency band. The original signal has frequency of $\pi/2$ instead of π according to the Nyquist's rule, after filtering the original signal by a highpass filter H and a lowpass filter L. In addition to the first level of 3D-DWT decomposition, by using LLL sub-band for another level of 3D-DWT, we can extract additional information from second level of 3D-DWT.
4. Introducing spectral band fusion by developing a single 2D image from a hyperspectral image cube by applying discrete wavelet transform (DWT) to each pixel along the spectral axis consecutively.
5. Introducing a new method to generate a 2D face image instead of 3D face cube by fusing spectral band after multiplying each band with a weight based on a

band-specific signal to noise ratio (SNR). Signal is assumed to dominate the noise resulting a high SNR in the middle band (visible band). Hence the middle band is used for the segmentation of the region of interest for SNR calculation. A mask with zero and one value is created in this band where one is given to each pixel in the homogeneous region and zero to all pixels in other clusters. Multiplying each spatial coordinate in each band with this mask help segment the homogeneous region in the specific band.

1.6 Thesis Overview

In Chapter 2 feature extraction, selection, dimensionality reduction and classification in state-of- the-art are discussed in detail for hyperspectral face recognition. Chapter 3 presents proposed method using 3D discrete wavelet transform for hyperspectral face recognition. It contains feature extraction, feature selection and classification for proposed method. Results are compared with several state-of-the-art facial hyperspectral images methods. In Chapter 4, we introduce a novel approach to improve hyperspectral face recognition using band-specific signal to noise ratio (SNR) based weighting to fuse spectral information acquired through different spectral bands into a single 2D image. Chapter 5 describes a new method to generate 2D face image instead of 3D face cube by utilizing DWT based fusion along spectral axis. Chapter 6 concludes this thesis and proposes future work based on this thesis work.

Chapter 2

LITERATURE REVIEW ON FACIAL HYPERSPECTRAL IMAGES

2.1 Introduction

Face recognition has a special role in biometrics in the past several years since it is non-contact process comparing to other biometrics systems such as fingerprint, palmprint and iris. This advantage helps capturing face images from a distance to identify the person without interacting [1]. Hence it has gotten special attention due to security difficulties by law enforcement.

Face recognition due to the variance in orientation, illumination and expressions is a challenging subject [3]. 2D face recognition has achieved significant development however still there are difficulties to overcome the mentioned challenges. One way to overcome mentioned challenges is the use of 3D face images which contain depth information [3]. However, 3D images also have limitations, but comparing to 2D face images they contain more information.

2.2 Hyperspectral Imagery

The human eye is able to recognize only a narrow interval of the light spectrum which is known as visible light (380nm-750nm). By considering spectrum beyond the visible light, the information that can be collected increase. Hence hyperspectral images (HSI) by collecting dozens of images in the narrow interval of energy wavelength (10-20 nm) are formed. Hyperspectral images mostly cover visible to near infrared

electromagnetic spectrum which has wavelengths between 400 nm to 1400 nm. Hyperspectral cameras capture images by using special techniques which combine spectral and spatial information. Hyperspectral image is a form of 3D images with two spatial and one spectral dimensions [30].

Hyperspectral imagery is a popular topic in remote sensing applications and due to the high cost of devices it was not popular topic in biometrics applications [58]. The recent improvements in technologies and optics made hyperspectral cameras more economical and accessible, hence methods utilized for remote sensing problems have been applied to biometrics applications.

2.3 Hyperspectral Face Recognition

Hyperspectral cameras operate in multiple narrow bands of visible spectrum and beyond to sample a face which resulting in more biometric information. This operation samples a face in spatial and spectral domain, resulting more information compared to traditional 2D grey/color images. This significant information can be related to distinct personal patterns originating from skin tissues, blood and organ structure [3]. There is another advantage regarding to spectral information which is distinguishing the real human face from mask or a photograph. In order to improve segmentation, classification and recognition of the images, hyperspectral face images based approaches employ spatial and spectral relationship simultaneously. Spectral dimension contains spectral information of faces which is related to inherent characteristics of subject. Since spectral dimension is captured across wide range of spectrum, it provides abundant information regarding to spectral response of each face which is different for each person [10].

By employing the application of facial hyperspectral image, difficulties encountered in visible light based face recognition systems, such as the variance in orientation, illumination or expressions can be minimized which helps to improve face recognition accuracy [3]. Besides these opportunities, hyperspectral images pose some challenges such as low signal to noise ratios, inter-band misalignment, high data dimensionality and data acquisition needs expensive cameras with multiple sampling in visible and near-infrared spectra. Despite of mentioned challenges, hyperspectral images represented in 3D-cubes are by far more capable in classification processes, which is also ideal for spoofing attacks. Bands near blue wavelength caused by high photon energy are usually the reason of low signal to noise ratio (SNR) [11]. Due to subject movements during hyperspectral face image acquisition inter-band misalignment occurs [11]. Fig 2.1 illustrates a hyperspectral face image that contains 33 bands from the PolyU hyperspectral face database (HSFD).

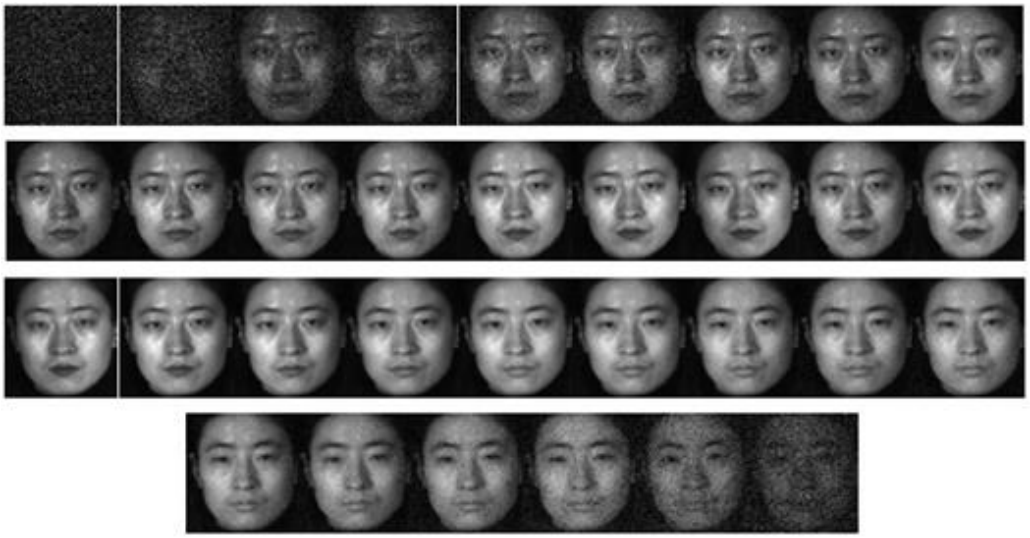


Figure 2.1: A sample of hyperspectral face cube of 33 bands from the Poly U-HSFD [4].

2.4 Feature Extraction

Due to high data dimensionality of hyperspectral face images feature extraction is a challenging task comparing to traditional grey/color images. This challenge is the main reason that there are not plenty of researches regarding to hyperspectral face recognition. Most researchers have extracted the features by sampling the hyperspectral face image [8-12-13] while some of them have just applied Principal Component Analysis (PCA) for feature extraction and dimensionality reduction [9].

2.4.1 2D Features Extraction Algorithms

In [12] Pan *et al*. used spectral features of hyperspectral face images in the near infrared range (NIR) (700 nm to 1000 nm). The spectral features from hair, forehead, cheeks, lips and chin for each subject are sampled manually for 31 bands. Mahalanobis distance is used to compare spectral features to perform face recognition accuracy rate.

Robila in [14] extracted spectral features of six regions (chin, nose, ear, eye, forehead and top lip) for eight subjects in the range of 400 nm to 900 nm. To comprehend the spectral power and spectral angle between each subject, means for each spectra location of six regions are computed. This process is applied to 120 bands of each hyperspectral face image.

Due to different physical absorption of face skin, Di *et al*. [9] used feature band selection to identify two feature band subsets and proposed three different methods for hyperspectral face recognition. These two feature band subsets are corresponded to hemoglobin compound which are located at 540 nm and 580 nm. They applied single band $(2\mathrm{D})^2\mathrm{PCA}$, whole bands $(2\mathrm{D})^2\mathrm{PCA}$ with decision fusion and band subset fusion-based $(2\mathrm{D})^2\mathrm{PCA}$ for hyperspectral face recognition.

Chen *et al.* [15] first perform denoising since hyperspectral face images contain significant amount of noise. Denoised face cube is cropped by eye coordinate, then log-polar transform is applied to each band of cropped face. To obtain invariant features to rotation and scale, 2D FFT is applied to log-polar face images to extract 2D Fourier Spectrum (FS).

In [16] Chen *et al.* extracted features for each cropped hyperspectral face cube with five different methods named local binary pattern (LBP), histogram of oriented gradients (HOG), log-polar transform, Gabor filter bank and Zernike moments. Then classifies each face into one of the existing classes. LBP [17] is the method which divide an image into local regions and label every pixel of an image by thresholding the local neighborhood pixels of each pixel with the center pixel and considers the result as a binary number. Histogram of oriented gradients (HOG) [18] describes object appearance and shape by the distribution of intensity gradients. HOG is computed for the pixels within small connected regions of the facial image then all these histograms will be concatenated. To implement HOG five steps, exist which are gradient computation, orientation binning, descriptor blocks, block normalization and a classifier [16]. Log-polar transform [19] has significant property which converts scaling factor in (x, y) coordinates to spatial shift in log-polar coordinates (ρ, θ). Fast Fourier transform (FFT) is applied to log-polar image to obtain Fourier spectrum which is scale/rotation invariant features along ρ. To convert Cartesian coordinates (x, y) to log-polar coordinates (ρ, θ) is formulated as,

$$\rho = log\sqrt{x^2 + y^2}\ and\ \theta = \tan^{-1}\frac{y}{x}\ if\ x > 0. \tag{2.1}$$

Gabor filter bank [20-21] by filtering a facial image with a bank of complex Gabor filter calculates the magnitude responses of an image. After downsampling magnitude

responses of an image will be normalized by using zero-mean and unit variance. After normalization process these magnitude responses are added to the output filtered image. The 2D Gabor filters in the spatial domain can be defined as a Gaussian kernel function by the following formula in the complex plane wave,

$$\psi_{u,v}(x,y) = \frac{f_u^2}{\pi\kappa\eta} e^{-\left(\left(\frac{f_u^2}{\kappa^2}\right)X^2+\left(\frac{f_u^2}{\eta^2}\right)Y^2\right)} e^{j2\pi f_u X} \quad (2.2)$$

where $X = xcos\theta_v + ysin\theta_v$ and $Y = -xsin\theta_v + ycos\theta_v$. The pixel coordinates are denoted by *x* and *y*, the frequency of the complex sinusoid and the orientation of the wavelet are denoted by f_u and θ_v respectively. The ratio between frequency of the complex sinusoid (center frequency) and the size of the Gaussian envelope is determined by κ and η. Zernike moments [22] are orthogonal moments which has the property of rotation invariant. It means the magnitudes of Zernike moments for any image after rotating does not change. Hence the features which are extracted from Zernike moments can easily be constructed to an arbitrary high order. The order which the reconstructed image is close to the original one is defined as the maximum order. Zernike moments features are only rotation invariant. Hence to obtain an image with scale/translation invariance, a normalization process using its regular moments is applied to the image. Then rotation invariant Zernike features can be extracted from normalized image. Zernike is defined as a set of complex polynomials which form a complete orthogonal set over the interior of the unit circle [23]. These polynomials are defined as,

$$V_{mn} = R_{nm}(\rho)e^{jm\theta} \quad (2.3)$$

where ρ defines the length of vector from origin to the pixel (x, y). $\theta = \tan^{-1}\frac{y}{x}$ is the angle between vector ρ and *x* axis. Radial polynomial ($R_{nm}(\rho)$) is defined as,

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2}(-1)^s \frac{(n-s)!}{s!(((n+|m|)/2)-s)!(((n-|m|)/2)-s)!}\rho^{n-2s} \quad (2.4)$$

The projection of the image *onto* orthogonal basis functions is defined by Zernike moments. The Zernike moment for a continuous image function $f(x, y)$ with order *n* and repetition *m* that just exists inside the unit circle ($x^2 + y^2 = 1$) is

$$A_{mn} = \frac{n+1}{\pi} \iint f(x, y)[V_{nm}(x, y)]^* \, dx \, dy \quad (2.5)$$

2.4.2 3D Features Extraction Algorithms

In [13] Shen and Zheng employed 3D Gabor wavelet to extract features form facial hyperspectral images. Gabor wavelets with different central frequencies has the advantage of analyzing and exploring the information in spatio-spectral domain in hyperspectral data cube simultaneously. To extract features from hyperspectral image cube a family of *M*×*N*×*Z* Gabor wavelets with different frequencies is selected which is defined as,

$$\Psi_{f_m,\varphi_n,\theta_k}(x, y, z), \;\; f_m = \frac{f_{max}}{2^m}, \; \varphi_n = \frac{n\pi}{N} \; , \theta_z = \frac{k\pi}{K} \quad (2.6)$$

The amplitude and orientations of central frequency are defined by f_m and (φ_n, θ_z) respectively. f_m is the frequency vector points with the same direction with different θ when $\varphi = 0$. To simplify the representation of the wavelets, $\Psi_{f_m,\varphi_n,\theta_z}$ denotes as $\Psi_{m,n,k}$. The information about local signal variances is represented by the inner product of signal with wavelet set $\Psi_{m,n,k}$ at location (x_c, y_c, z_c). The convolution results are considered to explore information at all possible locations in spatio-spectral domain *(x,y,z)*. The magnitude of convolution reveals the strength of variations across spatial and spectral domains.

Uzair *et al.* [11] employed 3D discrete cosine transform (DCT) to extract spatio-spectral features. 2D or 3D images can be expressed as a linear combination of

mutually uncorrelated cosine basis functions [24-25] by DCT. A compact energy spectrum of the signal can be generated by DCT. The low frequency coefficients of a compact energy spectrum are related to signal information. Hence the features are selected from low frequency coefficients. In [11] they extracted features from the whole hyperspectral face image by global 3D-DCT which has the advantage of modeling the spatio-spectral information simultaneously. They represented each hyperspectral facial cube by a small number of low frequencies DCT coefficients.

In [26] each hyperspectral face cube is divided into small overlapping 3D cubelets. Each of these 3D cubelets rearranged to the 2D matrix which each column corresponds to spectral response of all bands at a specific location. The first and second order statistics of each cubelets are computed by the mean vector and covariance of 2D matrix. In each cubelets the spread of information is indicated by the order statistics. The mean vector is added to each column of covariance matrix to mix the information of both order statistics. 3D cubelet is transformed to a singular variable by computing the Frobenius norm of mean and covariance matrix addition. Each 3D cubelets is replaced by its related single value resulting from Frobenius norm. Thus, by fusing information in all bands, a single 2D matrix represents hyperspectral face cube.

In [3] Vartak and Bharadi used Hybrid Wavelet Type I (HWI), Hybrid Wavelet Type II (HWII) and Kekre Wavelet (KW) to generate texture feature extraction. The PolyU Hyperspectral face database is used in this work. Hyperspectral face images contain 33 bands with front (F|), left (L) and right (R) side view for each subject. They grouped each hyperspectral face image into eleven sub-bands (each three bands consider as one band). In each hyperspectral face image three components (F, L and R) are considered which each of them is divided into 4×4 non-overlapping blocks for five level

decomposition. This process generates feature vectors by applying HWI, HWII and KW transforms to each subject. These feature vectors are applied to intra class and inter class testing to generate genuine and forgery classes. Multi-algorithmic and Multi-instance fusion are used to analyze these test results. For multi-algorithmic fusion the feature vectors of HWI, HWII and KW are fused and for multi-instance fusion the feature vectors of front, right and left side view samples are fused.

Liang *et al.* [10] proposed 3D Local Derivative Pattern (LDP) to analyze hyperspectral faces and encode each face cube into binary numbers as a 3D high order texture descriptor. To describe the changes in multi-directions and extract detailed features in multi-dimensional images, 3D directional derivative pattern and binarization function are employed in this method respectively. Since the spectral responses in hyperspectral images do not change roughly across the most wavelengths hence discriminative information related to spectral dimension may exist in some specific wavelengths. 3D LDP is a suitable method to extract features in hyperspectral face images as it functions in both spatial and spectral dimensions. After calculating features to represent an image as a vector histogram of a feature are estimated.

Chapter 3

3D-DWT BASED FEATURE EXTRACTION IN HYPERSPECTRAL FACIAL IMAGERY

3.1 Introduction

Recognition of humans by face, which is divided into two categories, i.e. global and local approaches has a special role in biometrics [12, 13]. Identifying individuals is one of the most basic examples of human face analysis [2]. Nowadays hyperspectral face recognition provides new opportunities for improving recognition accuracy since it contains more information compared to traditional 2D imagery which only contains spatial information (texture and structure). In facial hyperspectral images, samples are captured at multiple narrow bands within the neighborhood of visible light in the electromagnetic spectrum, which reveals significant information that is not evident in RGB images.

In hyperspectral images, the spectrum is divided into many bands, hence contains significant information regarding the skin at different wavelengths [4]. Having many bands, hyperspectral images contain extra significant information in spectral bands, compared to traditional grey/color facial image data [8, 9]. Hyperspectral imaging-based approaches employ spatial and spectral relationship simultaneously in order to further improve segmentation and classification of the images. Furthermore, distinct personal patterns originating from tissues, blood and organ structure can be captured using hyperspectral imaging. By employing the application of hyperspectral imaging,

difficulties encountered in visible light-based face recognition systems, such as the variance in orientation, illumination or expressions can be minimized [3].

In this chapter, three new methods for improving classification on hyperspectral face images are proposed. We employ 3D-DWT in each method to extract features from facial hyperspectral images. These methods are called 3D-subband energy (3D-SE), 3D-subband overlapping cube (3D-SOC) and 3D-Global Energy (3D-GE). 3D-DWT is employed to extract wavelet coefficients and, in each method, the energy vector is calculated from the wavelet coefficients in different manner. The extracted energy vector is regarded as the feature vector, where, the feature vectors are classified using *k* nearest neighbor (*k*-NN) and Collaborative Representation Classifier (CRC) classifiers. PolyU [4], CMU [5] and UWA [6] Hyperspectral Face Databases are used in these methods. Classification accuracies are evaluated by four test scenarios which 3D-GE method performance in terms of accuracy comparing to several existing methods improves significantly.

This chapter includes two contributions which are listed below:

1. The first contribution involves isolating the 3D data into frequency subbands using 3D Discrete Wavelet Transform (DWT). It is an appropriate procedure for feature extraction in 3D data since 3D-DWT decompose volumetric data in horizontal, vertical and spectral directions.
2. The second contribution is about feature extraction by using wavelet coefficients energy by three different approaches.

3.2 Databases

In order to provide a benchmark database for advance research, the PolyU-HSFD was developed by the Biometric Research Center (UGC/CRC) at Hong Kong Polytechnic University [4]. This database acquired using the CRI's VariSpec Liquid Crystal Tuneable Filter (LCTF). The database contains significant appearance (hair style and skin conditions), since it is constructed over a long period of time. The data cube contains 33 bands in the wavelength between 400 nm and 720 nm by step size of 10 nm. The datacube size is 220×180×33. Each subject has three different types of views, frontal (F), right (R) and left (L). There are 48 subjects in the database which contains 151, 125 and 124 datacubes for front, right and left images, respectively. Fig. 3.1 shows 33 bands of a face cube with frontal view in PolyU hyperspectral face database (HSFD).

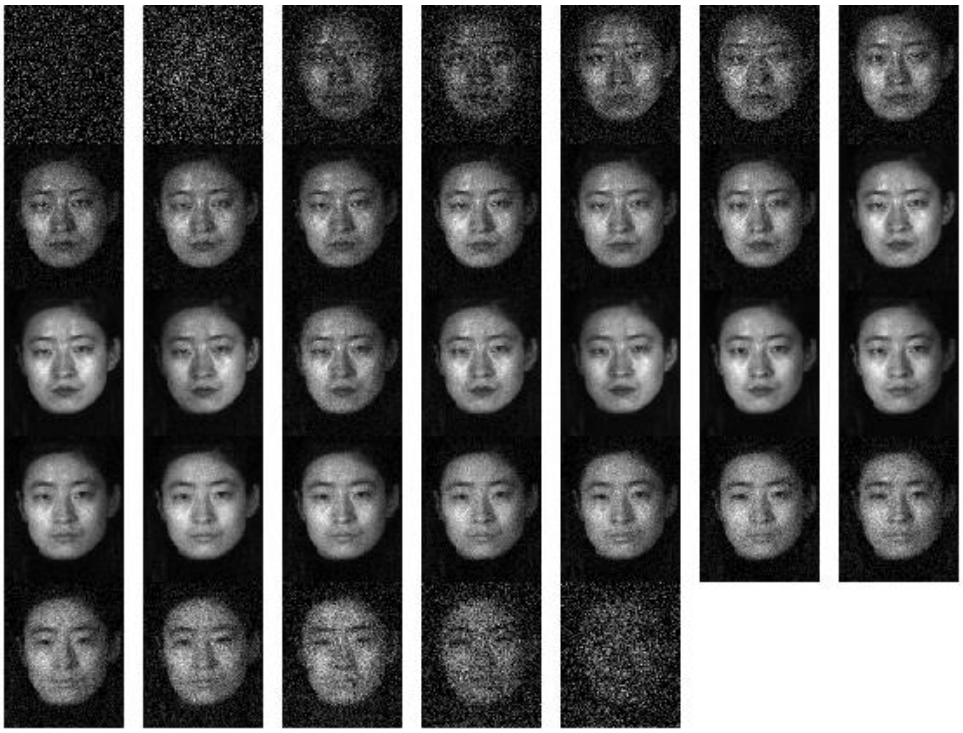


Figure 3.1: Frontal view of a hyperspectral face cube containing 33 bands form PolyU-HSFD.

The Robotics Institute of Carnegie Mellon University developed CMU hyperspectral face database [5]. The prototype spectro-polarimetric camera is used to develop this database. There are 48 subjects. Each datacube in this database contains 65 bands in range of 450-1090 nm with a step size of 10 nm. The database contains different sessions for each person according to lighting combinations and each person has 4-20 cubes according the light combinations. In our experiments the sessions with all lights on and all subjects which have 1-5 cubes are chosen. Fig. 3.2 illustrates even bands of a face cube in CMU hyperspectral face database (HSFD).

Figure 3.2: A sample of hyperspectral face cube containing even bands form CMU-HSFD.

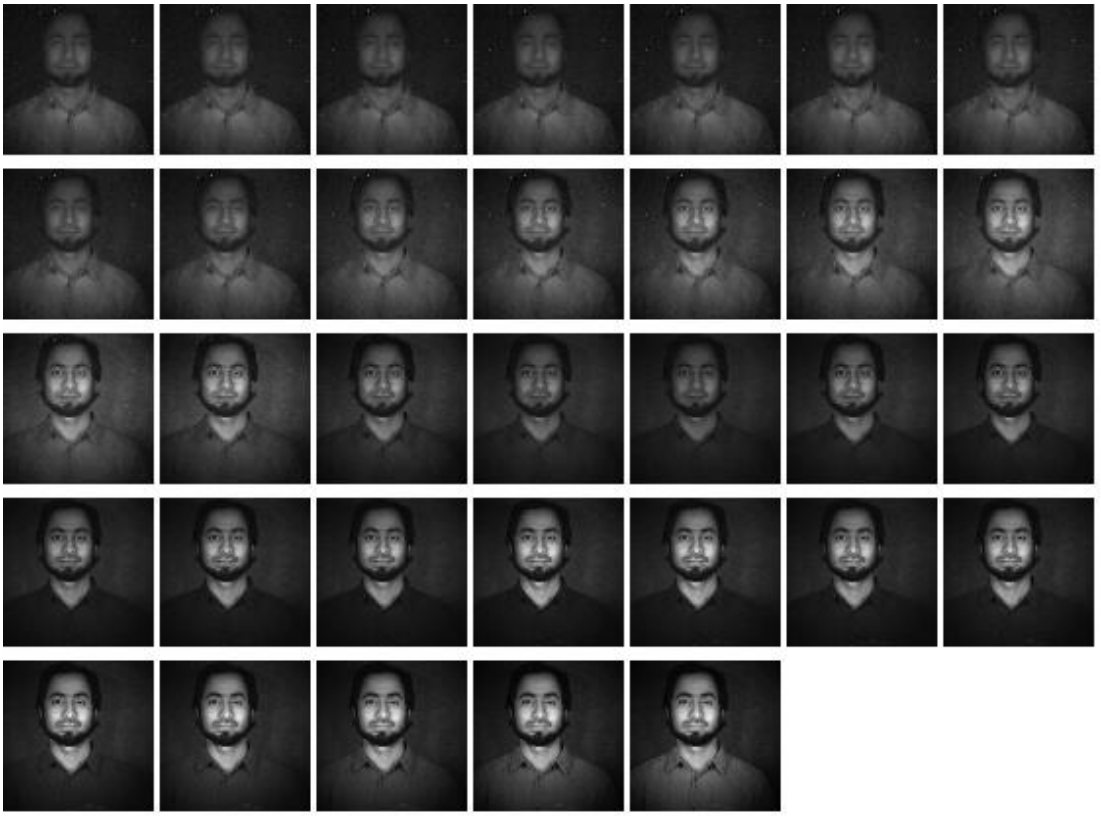


Figure 3.3: A sample 33 bands of hyperspectral face cube form UWA-HSFD.

The UWA [6] hyperspectral face database which is developed by an indoor imaging system using a CRI's VariSpec LCTF integrated with a photon focus camera. There are 70 subjects in the database which contains 120 hyperspectral cubes. Each data cube contains 33 bands in the wavelength between 400 nm and 720 nm by step size of 10 nm. The datacube size is 1024×1024×33. In Fig. 3.3 a sample of face cube containing 33 bands is shown from UWA hyperspectral face database (HSFD).

3.3 Three Dimensional Discrete Wavelet Transform

In 3D discrete wavelet transform (DWT), the 1D analysis filter bank is applied to each of the three dimensions. Hence the procedure can be considered as a combination of three 1D-DWT in the X, Y (spatial) and Z (depth) dimensions [27-28]. The 3D-DWT which isolates the data into frequency sub-bands can be regarded as a more advanced preprocessing method for 3D coding compared with 2D methods [29]. It considers the

correlation of 3D data cubes, which helps to improve the compression. The basic idea is to represent a signal as a superposition of wavelets. 3D-DWT has an advantage by decomposing volumetric data in horizontal, vertical and depth directions unlike 2D-DWT which decomposes only in horizontal and vertical dimensions [30]. For the hyperspectral images, 3D-DWT is performed by applying one dimensional DWT filter banks on three spatio-spectral dimensions [31]. Data cube of size $B_1 \times B_2 \times B_3$ after applying 1D-DWT to the first, second and third directions, result two, four and eight sub-bands, each of size $(\frac{B_1}{2} \times B_2 \times B_3)$,$(\frac{B_1}{2} \times \frac{B_2}{2} \times B_3)$,$(\frac{B_1}{2} \times \frac{B_2}{2} \times \frac{B_3}{2})$ respectively. Fig. 3.4 demonstrates 3D data decomposition into eight sub-bands after applying single level decomposition.

The 3-D DWT centered by a tensor product [32],

$$I^{(x,y,z)} = (L^x \oplus H^x) \otimes (L^y \oplus H^y) \otimes (L^z \oplus H^z)$$
$$= L^x L^y L^z \oplus L^x L^y H^z \oplus L^x H^y L^z \oplus L^x H^y H^z \oplus H^x L^y L^z \oplus H^x L^y H^z \oplus H^x H^y L^z \quad (3.1)$$
$$\oplus H^x H^y H^z$$

where the direct sum and tensor product are expressed by $\oplus$ and $\otimes$, respectively. The spatial (horizontal and vertical) domains and the spectral dimension of an image are represented by X, Y and Z directions respectively. Along three dimensions, the parameters *L* and *H* represent the low-pass and high-pass filters respectively.

The subband that passed through the low pass filter in horizontal, vertical and depth directions is called *LLL*. This band refers to the approximation of data cube. Marginal plane in each direction is shown by subbands which are passed through high-pass filters in only one direction. The boundary lines are defined by subbands which are passed through high-pass filters in two directions. Vertexes angle of the data cube is

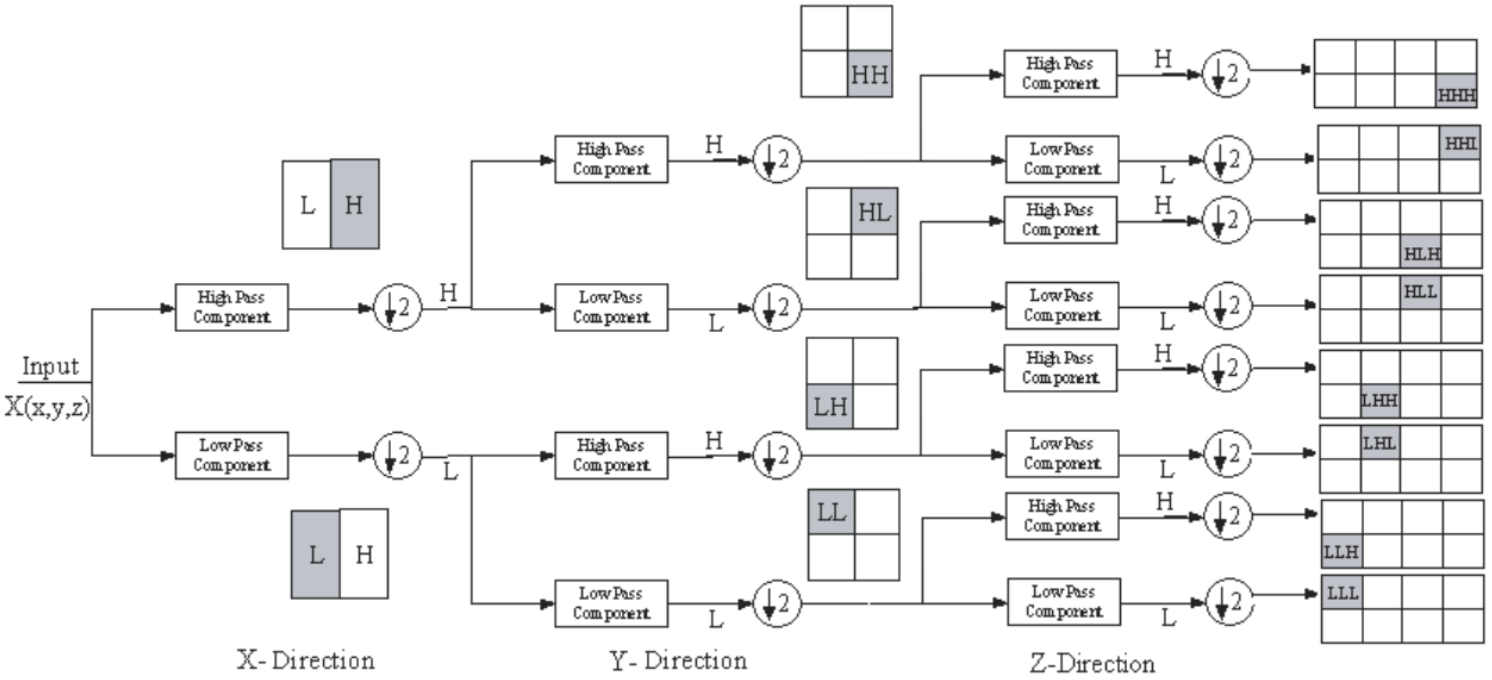


Figure 3.4: One level 3D-DWT procedure [33].

shown by *HHH* band which passed through high pass filter in horizontal, vertical and spectral directions [30].

3.4 The State-Of-The-Art Hyperspectral Feature Extraction

3D-DWT has been used efficiently in video coding [32], 3D medical analysis [34] and multi band remotely sensed images. Xian Guo *et al.* [35] employed 3D-DWT to extract features by applying it to the series of local cubes around the central pixel. By implementing 3D-DWT along spatio-spectral dimensions, wavelet coefficients from eight subbands capture the variations in the respective dimensions. The local 3D-DWT texture is represented by the energy values of the wavelet coefficients which are defined in three ways: pixel-based shift (PDWT), non-overlapping (WDWT), and overlapping cube (OWDWT), respectively.

3.4.1 Pixel-Based Shift

In this method, a local cube around each voxel (3D pixel) is used to process 3D texture and then the local texture measure is calculated by the quadratic sum of the wavelet coefficients in respective subbands. $E(W)$ is the subband energy where W is a local cube with B as horizontal, vertical and N as spectral dimensions respectively. $P(i,j,k)$ represents the wavelet coefficient in the local cube centered by the voxel (i,j,k).

$$E(W) = \sum_{i=1}^{B/2^L} \sum_{j=1}^{B/2^L} \sum_{k=1}^{N/2^L} P(i,j,k)^2 \tag{3.2}$$

Each pixel in the given image is used as the center of a local moving cube in PDWT which is subject to a large computational cost.

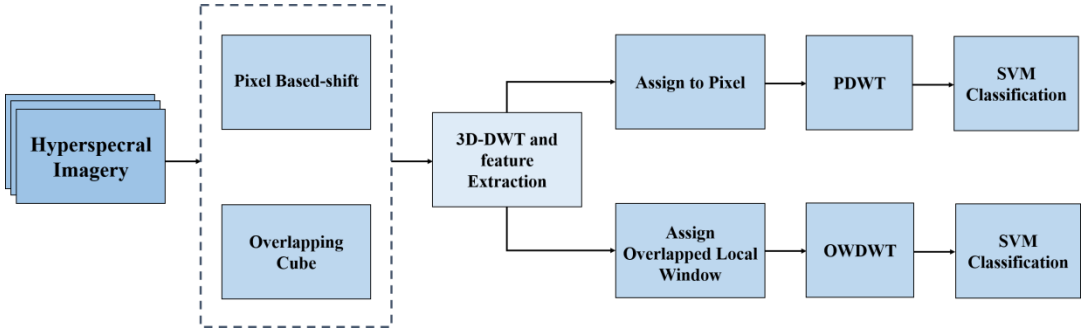


Figure 3.5: 3D- DWT framework for hyperspectral feature extraction followed by classification [35].

3.4.2 Overlapping Cube

In this section, there exist four adjacent cubes, i.e., surrounding W_1, W_2, W_3, and W_4 cubes and the overlapping central cube. To avoid edge-blurring effects and spatial information loss, Guo *et al.* [35] consider the spatial relationship of these adjacent cubes. Weighted feature representation is introduced to deal with the pixels in the overlapping area. At each local cube the 3D wavelet features are extracted to fuse spatio-spectral information in overlapping cubes, where the overlapping 3D texture is defined as a weighted linear combination of the neighboring texture energy values.

$$OWDWT = \frac{\omega_1 E(W_1) + \omega_2 E(W_2) + \cdots + \omega_n E(W_n)}{\omega_1 + \omega_2 + \cdots + \omega_n} \tag{3.3}$$

The weight and texture measurement of local cube W_n are defined by ω_n and $E(W_n)$, respectively. The weight is defined by the Euclidean distance between a local cube and the overlapping central cube. Fig. 3.5 demonstrates the 3D-DWT feature extraction framework for hyperspectral image classification [35].

3.5 Proposed Hyperspectral Face Feature Extraction Methods

Considering, PDWT and OWDWT locally extracting the 3D texture [35] for classification in hyperspectral satellite imagery, three alternative global methods are proposed in hyperspectral face recognition for improved performance. The proposed

methods are 3D-subband energy (3D-SE), 3D-subband overlapping cube (3D-SOC) and 3D-Global Energy (3D-GE).

3.5.1 3D-Subband Energy (3D-SE)

In this section, 3D-DWT feature extraction based on extracting the wavelet coefficients of the whole cube is proposed. The texture property is defined by the subband cube around each wavelet coefficient.

$$E_g(W) = \sum_{i=1}^{M/2} \sum_{j=1}^{N/2} \sum_{k=1}^{Q/2} P_g(i,j,k)^2 \tag{3.4}$$

$$\varepsilon = \left[E_g(LLL)\,, E_g(LLH), \ldots\ldots, E_g(HHL)\,, E_g(HHH)\right] \tag{3.5}$$

The energy of the whole cube is defined by $E_g(W)$, where W represents the 3D-DWT subband cube with M, N and Q in horizontal, vertical and spectral dimensions respectively. The wavelet coefficient in the subband cube centered by the voxel (i,j,k) is defined as $P_g(i,j,k)$. The ε is the energy vector of all 8 subbands. The framework for 3D-SE is shown in Fig. 3.6.

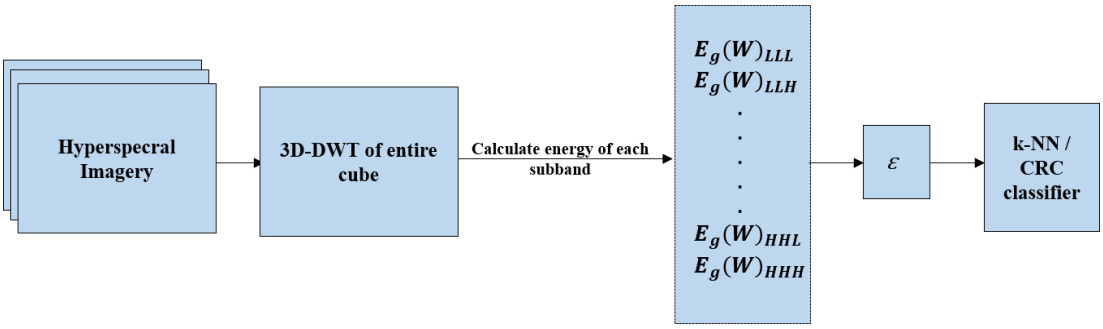


Figure 3.6: 3D-DWT framework for hyperspectral feature extraction using 3D-SE.

3.5.2 3D-Subband Overlapping Cube (3D-SOC)

In this section, there exist four adjacent cubes, i.e., surrounding W_1, W_2, W_3, and W_4 cubes and the overlapping central cube. To avoid edge-blurring effects and spatial information loss, the spatial relationship of these adjacent cubes is considered.

Weighted feature representation is introduced to deal with the pixels in the overlapping area. At each local cube, the 3D wavelet features are computed globally to fuse spatio-spectral information in overlapping cubes, where the overlapping 3D texture is defined as a weighted linear combination of the neighboring texture energy values.

$$\Omega_g = \frac{\omega_1 E_g(W_1) + \omega_2 E_g(W_2) + \cdots + \omega_4 E_g(W_4)}{\omega_1 + \omega_2 + \omega_3 + \omega_4} \tag{3.6}$$

$$\Phi = [\Omega_g(\mathrm{LLL}), \Omega_g(\mathrm{LLH}), \ldots\ldots, \Omega_g(\mathrm{HHL}), \Omega_g(\mathrm{HHH})] \tag{3.7}$$

In (3.6), the weighted linear combination of the neighboring global texture measures is defined by Ω_g. The weighted linear combination of all eight subbands is represented as vector Φ. Fig. 3.7 illustrates the 3D-DWT framework using 3D-SOC for hyperspectral feature extraction.

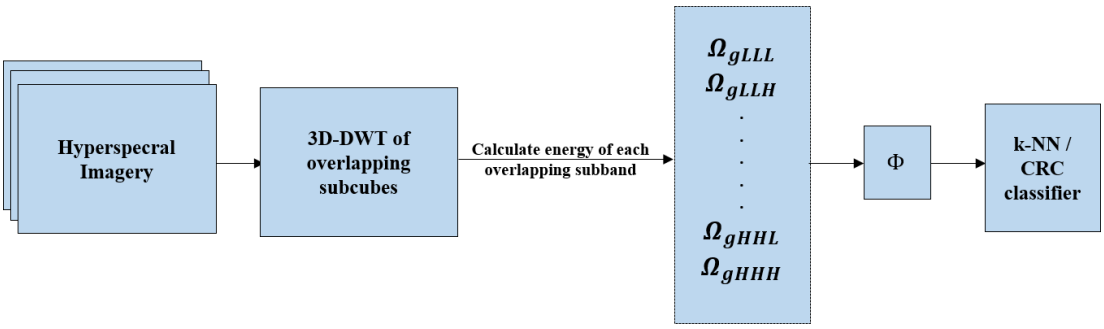


Figure 3.7: 3D-DWT framework for hyperspectral feature extraction using 3D-SOC.

3.5.3 3D-Global Energy (3D-GE)

In this section, the 3D wavelet coefficients of the whole cube after the first and second level of 3D-DWT are extracted. The second level 3D-DWT is extracted from LLL subband of considered and the observation was that the performance saturates beyond two levels of 3D-DWT transform. The wavelet coefficients characterize the texture feature vectors given in equations (3.9) and (3.11).

$$E_G^{(1)}(Z) = \sum_{i=1}^{M/2} \sum_{j=1}^{N/2} \sum_{k=1}^{Q/2} P_G^{(1)}(i,j,k)^2 \tag{3.8}$$

$$\varepsilon_{\mathrm{G}}^{(1)}=[\mathrm{E}_{\mathrm{G}}^{(1)}(1)^{\mathrm{LLL}},\mathrm{E}_{\mathrm{G}}^{(1)}(2)^{\mathrm{LLH}},.....,\mathrm{E}_{\mathrm{G}}^{(1)}(7)^{\mathrm{HHL}},\mathrm{E}_{\mathrm{G}}^{(1)}(8)^{\mathrm{HHH}}] \tag{3.9}$$

$$E_{GLLL}^{(2)}(Z) = \sum_{i=1}^{M/4} \sum_{j=1}^{N/4} \sum_{k=1}^{Q/4} P_{GLLL}^{(2)}(i,j,k)^2 \tag{3.10}$$

$$\varepsilon_{GLLL}^{(2)}=[\mathrm{E}_{GLLL}^{(2)}(\mathrm{LLL}),\mathrm{E}_{GLLL}^{(2)}(\mathrm{LLH}),..., \mathrm{E}_{GLLL}^{(2)}(\mathrm{HHL}),\mathrm{E}_{GLLL}^{(2)}(\mathrm{HHH})] \tag{3.11}$$

$E_G^{(1)}(Z)$ and $E_{GLLL}^{(2)}(Z)$are the first and second level energy of global cube, respectively. In equation 3.10, *Z* represents the global cube with *M*, *N* and *Q* dimensions, which correspond horizontal, vertical and spectral dimensions respectively. $P_G^{(1)}(i,j,k)$ and $P_{GLLL}^{(2)}(i,j,k)$ are the first and second level wavelet coefficients corresponding to the voxel (i,j,k) in the global cube respectively. $\varepsilon_G^{(1)}$ and $\varepsilon_{GLLL}^{(2)}$ are the energy vectors of the first and second level of all eight subbands. $\varepsilon_G^{(1)}$ and $\varepsilon_{GLLL}^{(2)}$ are concatenated to make a vector, $\varepsilon_G^T$with 16 coefficients. The 3D-DWT based 3D-GE method is demonstrated in Fig. 3.8. First and second level 3D-DWT of LLL subband for a hyperspectral face cube for Poly-U, CMU and UWA hyperspectral face databases are demonstrated in Fig. 3.9, Fig. 3.10 and Fig. 3.11 respectively.

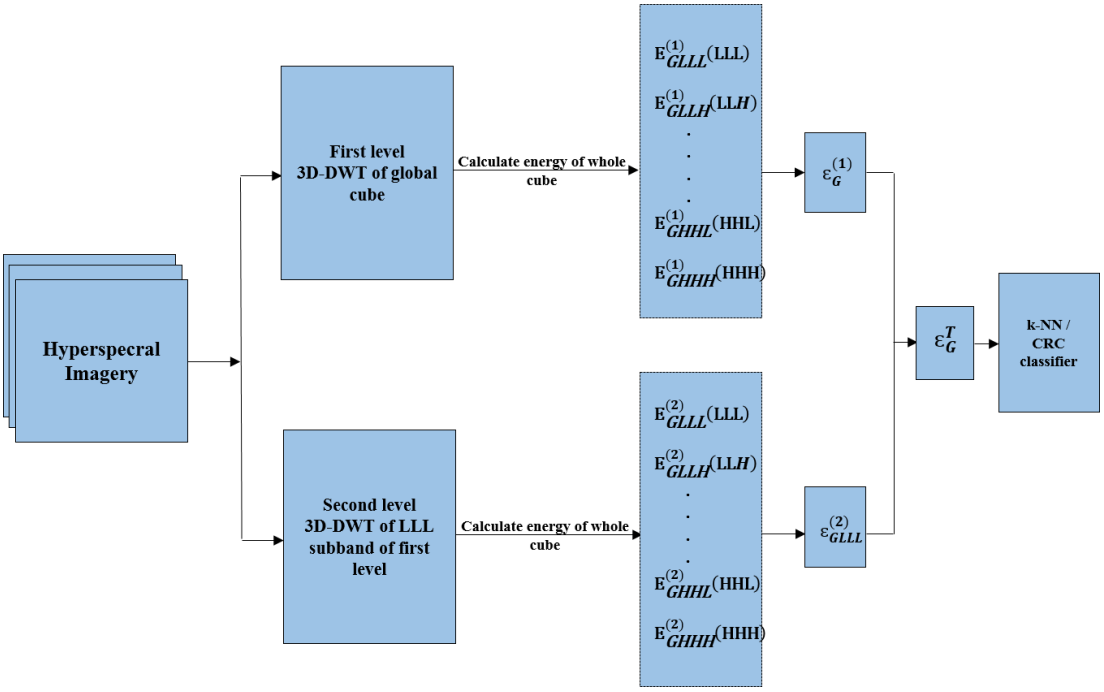
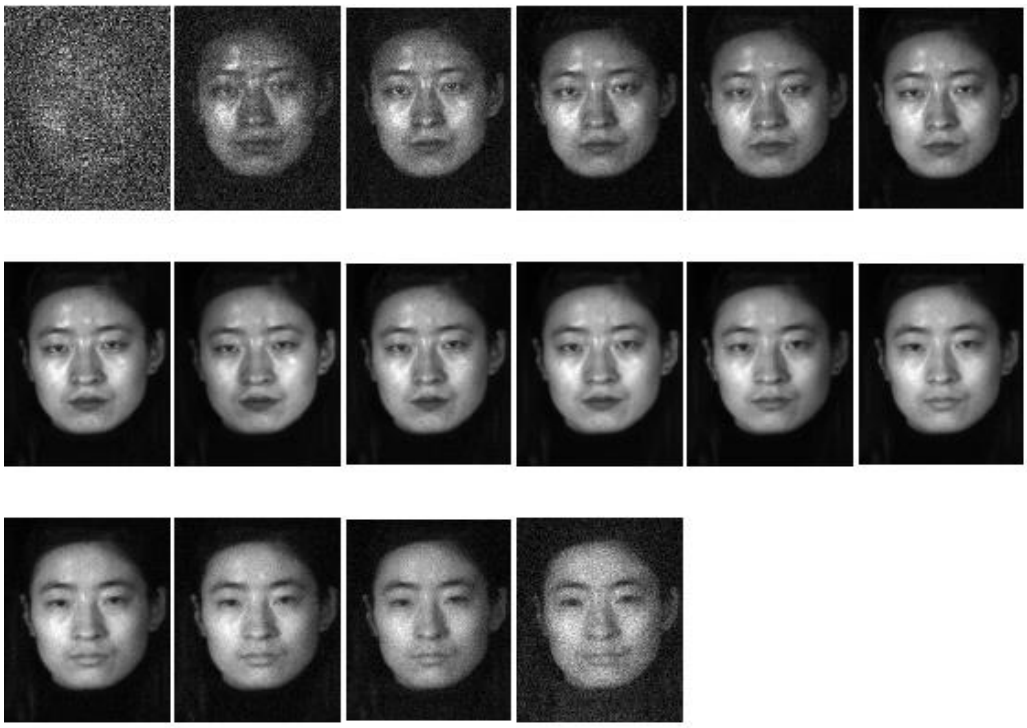
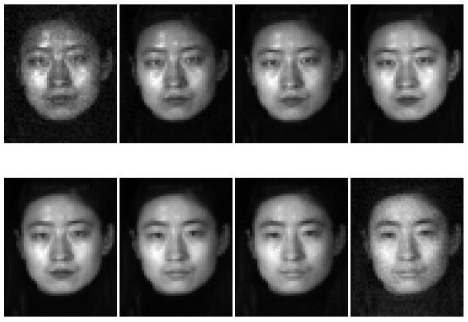


Figure 3.8: 3D-DWT framework for hyperspectral feature extraction using 3D-GE.



(a)



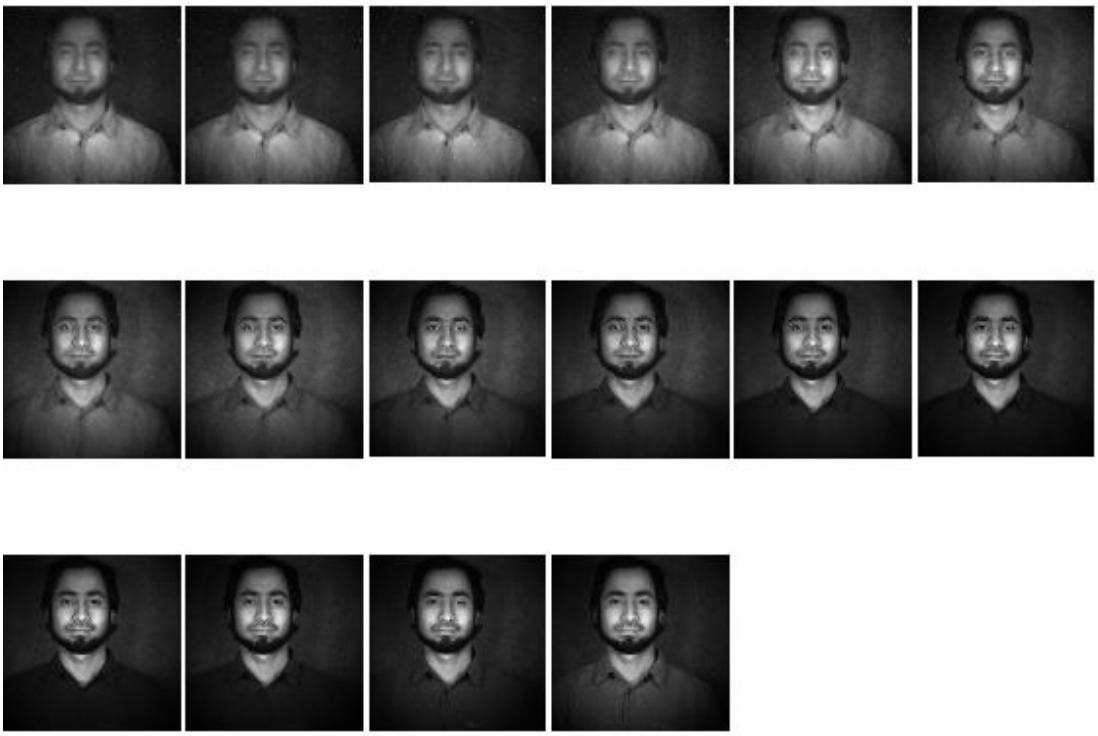
(b)

Figure 3.9: First and second level 3D-DWT of LLL subband for a hyperspectral face cube from PolyU-HSFD. (a) first level, (b) second level.

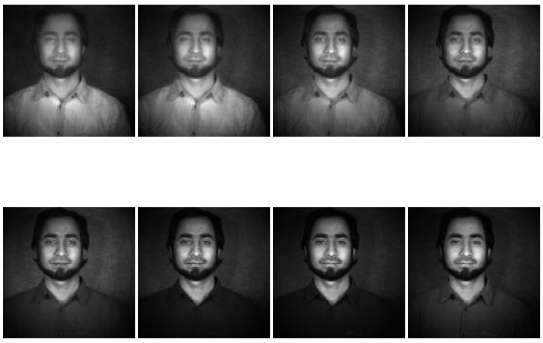
(a)

(b)

Figure 3.10: First and second level 3D-DWT of LLL subband for a hyperspectral face cube from CMU-HSFD. (a) first level, (b) second level.



(a)



(b)

Figure 3.11: First and second level 3D-DWT of LLL subband for a hyperspectral face cube from UWA-HSFD. (a) first level, (b) second level.

3.6 Classifiers

The *k*-NN classifier is used for evaluating the recognition rate in terms of accuracies in this thesis. Additionally, in this chapter, CRC (collaborative representation classifier) is employed for recognition rate comparison with *k*-NN classifier. In the following subsection, *k*-NN and CRC classifiers and their techniques are described.

3.6.1 *k*- Nearest Neighbor Classifier (*k*-NN)

k-Nearest Neighbor (*k*-NN) is one of the common classifiers that is used in face recognition. Image is classified by a majority vote that is given by the *k* neighbours of it [38]. A distance matrix which is the Euclidean distance between the testing image feature and each training image feature is calculated. The first *k* elements in summation value of distance matrix which are ordered in ascending manner are selected to choose the majority class value for image classification. Each sample belongs to a known class (C_n) and the test image is categorized to the class which has the majority value according to the first *k* elements [30-40].

3.6.2 Collaborative Representation Classifier (CRC)

Let define a matrix $B = [B_1, B_2, \ldots., B_M]$ whose columns are the training image feature vectors, where *M* is the number of subjects in the database, and the training samples of the subject is defined by $B_k = [B_{k1}, B_{k2}, \ldots., B_{kC}]$, where $k \in [1, M]$, and *C* is the number of bands in the hyperspectral face image. Let the column of $A = [a_1, a_2, \ldots., a_C]$ be a testing hyperspectral face cube with *C* faces.

The optimization problem in Equation 3.12 corresponds to Collaborative representation classifier (CRC) [41],

$$\tilde{\sigma}_k = argmin_\sigma \|a_k - B\sigma\|_2^2 + \lambda\|\sigma\|_2^2 \qquad (3.12)$$

where λ is a regularization parameter and $\tilde{\sigma}_k = [\tilde{\sigma}_{k1}, \tilde{\sigma}_{k2}, \ldots, \tilde{\sigma}_{kM}]$. The solution for Equation 3.12 can be easily derived as:

$$\tilde{\sigma}_k = (B^T B + \lambda I)^{-1} B^T a_k \qquad k \in [1, C] \qquad (3.13)$$

$P = (B^T B + \lambda I)^{-1} B^T$ and it does not depend on a_k hence it can be pre-calculated as a projection matrix just simply project a_k onto *P* via $\boldsymbol{Pa_k}$ [41]. The CRC classifies a face a_k as:

$$identity(a_k) = \ argmin_i \left\{ \frac{\|a_k - B_i \tilde{\sigma}_{ki}\|_2}{\|\tilde{\sigma}_i\|_2} \right\}, \qquad i \in [1, M] \qquad (3.14)$$

CRC classifies each testing image to the class with minimal regularized reconstruction error.

3.7 Experimental Results

In this section we first clarify four different test scenarios which are utilized for a fair comparison, then the recognition rates of proposed methods are compared with state-of-the-art hyperspectral face image recognition methods.

3.7.1 Test Scenarios Distribution

1. First scenario: First 38 subjects with all data cubes (hyperspectral image) of frontal hyperspectral face images with 33 and 24 bands in PolyU-HSFD are selected. Due to the high noise, first six and the last three bands are removed [13, 26] leaving behind 24 bands. The leave-one-out method is applied to train and test the classifier.
2. Second scenario: In this scenario, the first 25 subjects of frontal hyperspectral face images with 4 data cubes including 33 and 24 bands in PolyU-HSFD were chosen. Two of four cubes for each subject were selected randomly for training. The rest were selected as the test sets [11].
3. Third scenario: All data cubes of the first 25 subjects (frontal, right and left) of hyperspectral face images with all the bands in PolyU-HSFD were selected. Leave one out validation approach is adopted in the training and testing of the classifiers.
4. Fourth scenario: In this scenario, first 25 subjects of frontal face images with 24 spectral bands in PolyU-HSFD are used in the experiment. For each subject, 2 cubes are selected randomly for training. The rest of the 63 cubes are used for testing [1, 11]. In order to verify our results in other databases, we used

CMU-HSFD and UWA-HSFD in this scenario. In CMU, we selected all subjects which have 1-5 cubes and for each subject one cube is randomly used for the training set and the rest of the cubes as a testing set [11]. In UWA, for each subject, one cube is randomly selected for training and the rest 50 cubes as a testing set [11].

3.7.2 Experimental Results and Comparison

The performance analysis involved using the average accuracy and standard deviation (STD) through the process of randomly picking training and testing datasets ten times for all scenarios. The accuracy is defined as the percentage of the ratio of the correctly classified hyperspectral faces over the total number of hyperspectral faces in the test set. We set the number of nearest neighbor predictor three in the *k*-NN classifier and the regularization parameter λ to 0.001 in CRC classifier.

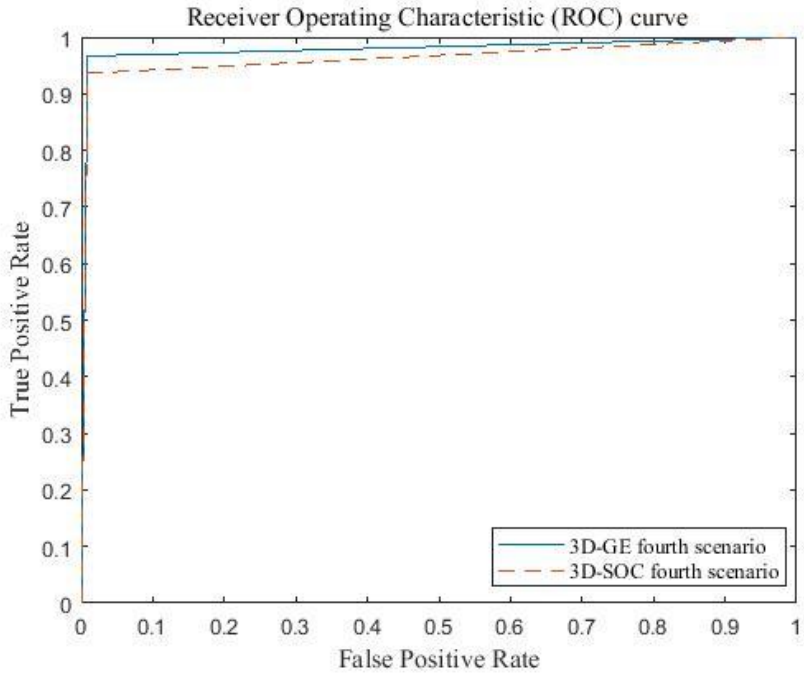


Figure 3.12: ROC curve of 3D-GE and 3D-SOC for a subject from the PolyU-HSFD.

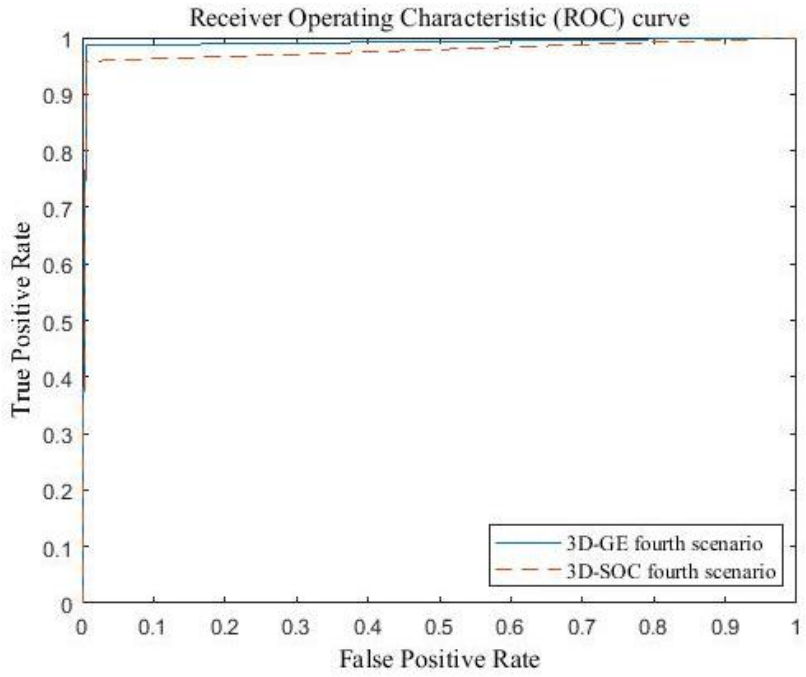


Figure 3.13: ROC curve of 3D-GE and 3D-SOC for a subject from the CMU-HSFD.

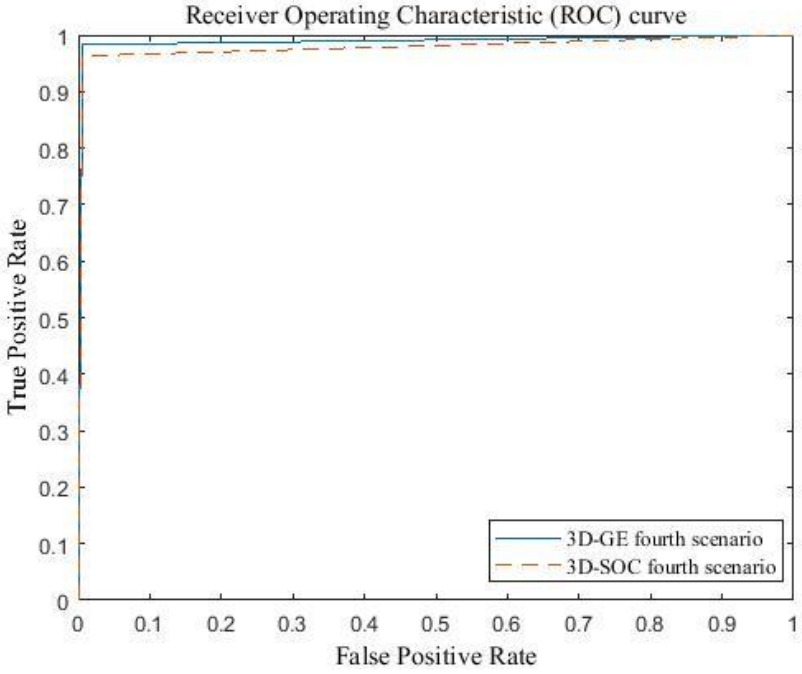


Figure 3.14: ROC curve of 3D-GE and 3D-SOC for a subject from the UWA-HSFD.

Fig. 3.12, Fig. 3.13 and Fig. 3.14 show the receiver operating characteristic (ROC) curves for 3D-GE and 3D-SOC methods under fourth scenario in PolyU-HSFD, CMU-HSFD and UWA-HSFD, respectively. The ROC curves show the true positive rate (TPR) against false positive rate (FPR).

$$TPR = \frac{TP}{TP + FN} \tag{3.15}$$

$$FPR = \frac{FP}{FP + TN} \tag{3.16}$$

$$ACC = \frac{FP + TN}{TP + TN + FP + FN} \tag{3.17}$$

where *TP*, *FN*, *FP* and *TN* denote true positive, false negative, false positive and true negative which are between interval [0,1]. The area under the curve (AUC) indicates the classifier performance, the larger AUC implies the higher classification performance. Since the ROC curve in multiclass problems require $n^2 - n - 1$ dimensional hypersurface, hence the visualization of it is impossible. However, it is possible to extend two-class ROC curve to multiclass [42]. This approach has two forms: *one versus all* and *one versus one*. In *one-versus-all* form, the classification performance is evaluated for class *'one'* and all of the remaining classes are considered into a single class *'not one'*. Therefore, details of specific misclassification errors of the class *'not one'* are not available [16]. In this context, we have adopted *one-versus-all* approach.

Table 3.1 shows that the experimental results after 3D-SOC, 3D-SE and 3D-GE are applied to first test scenario. The proposed 3D-GE has achieved the recognition with 94.73% ± 2.16 and 95.57% ± 1.41 accuracies in the first scenario by using *k*-NN and CRC classifiers, respectively. The experimental results of second testing scenario are

shown in Table 3.2. The proposed 3D-GE for 33 and 24 bands by using [*k*-NN/CRC] classifiers has been tested and [89.33% ± 2.32 / 91.76% ± 1.57] and [86.71% ± 1.11 / 90.36% ± 1.15] accuracies have been obtained, respectively. The results in Tables 3.1 and 3.2 show that using all bands (33) instead of 24 bands improves the recognition rate. The results suggest that the removed noisy bands contain useful information and 3D-DWT has an advantage of extracting them. Table 3.3 shows significant improvement of recognition over [3], after the proposed 3D-GE is applied to third scenario for frontal, right and left 94.16% ± 1.23, 78.94% ± 2.19 and 83.2% ± 1.67 by using *k*-NN classifier and 94.43% ± 2.1, 83.35% ± 2.3 and 87.97% ± 1.6 by using CRC. The experimental results of fourth testing scenario are shown in Table 3.4. The proposed 3D-GE has achieved the recognition 96.66% ± 1.2 and 98.61 ± 1.3 and 98.28% ± 1.05 by using CRC classifiers in PolyU-HSFD, CMU-HSFD and UWA-HSFD, respectively.

Table 3.1: First scenario recognition accuracy (%)

Bands	**PDWT [35]**	**OWDWT [35]**	**3D-SE**	**3D-SOC**	**Proposed 3D-GE (k-NN)**	**Proposed 3D-GE (CRC)**
33	94.02±1.15	94.13±1.84	94.28±2.50	94.35±2.06	94.73±2.16	**95.57±1.41**
24	83.43±2.10	83.95±1.67	89.58±2.43	89.91±1.32	90.52±2.53	**92.34±1.15**

Table 3.2: Second scenario recognition accuracy (%)

Bands	BS-WFD [9]	Best Individual 3D Gabor [13]	3D-SE	3D-SOC	Proposed 3D-GE (*k*-NN)	Proposed 3D-GE (CRC)
33	-	-	87.66±1.70	88.67±1.53	89.33±2.32	**91.76±1.57**
24	79.00	82.00	85.33±2.00	85.45±2.50	86.71±1.11	**90.36±1.15**

Table 3.1 shows that the recognition accuracy for proposed method is lower if all bands do not be used in feature extraction. Comparing Table 3.1 and Table 3.2 shows that adopting 25 subjects instead of 38 (all subjects) subjects in feature extraction help improving the recognition rate regarding more subjects in training set. CRC classifier comparing to k-NN has higher recognition accuracy since CRC classifies each testing image to the class with minimal regularized reconstruction error.

Table 3.3: Third scenario recognition accuracy (%)

Type of views	HWII [3]	Best Individual 3D Gabor [13]	3D-SE	3D-SOC	Proposed 3D-GE (*k*-NN)	Proposed 3D-GE (CRC)
Front	74.3	93.3±2.42	94±1.45	94.16±1.23	94.16±1.23	**94.43±2.1**
Right	77.5	74.45±1.81	77.61±1.9	78.94±2.19	78.94±2.19	**83.35±2.3**
Left	74.3	79.85±2.32	81.62±1.16	83.2±1.67	83.2±1.67	**87.97±1.6**

In Table 3.3 front view for proposed method has higher recognition accuracy comparing o right and left views. This occur because of more features are exist in whole face comparing to just right or left side of a face. CRC classifier comparing to k-NN has higher recognition accuracy since CRC classifies each testing image to the class with minimal regularized reconstruction error.

Table 3.4: Fourth scenario recognition accuracy (%)

Databases	Gabor wavelet [15,26]	3D-SOC	Log-polar FFT2 [15]	3D-LDP [10]	Band fusion +PLS [26]	Feature extraction and CRC [16]	Proposed 3D-GE (CRC)
PolyU	91.3±2.1	93.61±2.2	94.6±2.5	95.3±1.6	95.2±1.6	96.4±2.3	**96.66±1.2**
CMU	91.6±2.9	95.83±1.8	95.6±1.7	94.8±2.6	**99.1±0.6**	98.0±0.7	98.61±1.3
UWA	91.5±3.07	96.13±2.1	-	-	98.2±1.2	-	**98.28±1.05**

In this contribution 3D-DWT is employed for feature extraction which has a computational complexity of $\mathcal{O}(N^3)$ where $N$is an input's dimension. It is same/high comparing to mentioned state of the art methods in Table 3.4. In [13], they employed 3D-Gabor wavelet to extract features of hyperspectral face images which has the same computational complexity of $\mathcal{O}(N^3)$. In [15-16], features are extracted by 2D Fast Fourier Transform (FFT) from the log-polar images hence the computational complexity of FFT is $\mathcal{O}(NlogN)$. In [26], the first and second order statistics of each cubelets (section 2.4.2) are computed by the mean vector and covariance of 2D matrix which covariance matrix computation has a computational complexity of $\mathcal{O}(N^2)$.

The proposed 3D-GE improves the recognition rates since in the second-level 3D-DWT the extracted features from LLL subband of first-level 3D-DWT contain more information.

3.8 Conclusion

In this chapter three alternative 3D-DWT-based methods for the feature extraction for hyperspectral face recognition and classification, namely 3D-subband energy, 3D-subband overlapping cube and 3D-global energy are proposed. Hyperspectral faces are represented by the energy measures calculated from the respective subbands using the first and second level of 3D-DWT. The extracted feature vectors were used as the input to the *k*-NN and CRC classifiers for performance measurement. Experimental results on PolyU-HSFD, CMU-HSFD and UWA-HSFD using the proposed methods were compared with several alternative methods on hyperspectral face recognition. The results show that the 3D-GE performance in terms of accuracy outperforms several existing methods using PolyU, CMU and UWA databases.

Chapter 4

SNR BASED FUSION OF SPECTRAL BANDS FOR IMPROVED HYPERSPECTRAL FACE RECOGNITION

4.1 Introduction

In this chapter a new method is proposed to improve hyperspectral face recognition. High dimensionality and low signal to noise ratio are the challenges which are posed by hyperspectral images. Low signal to noise ratio appears in certain spectral bands which are mostly located near the blue wavelength. We proposed a new method to overcome these problems without losing information from spectral bands.

This chapter is included two contributions which are listed below:

1. In the first contribution, the middle band is chosen as the reference for determining homogeneous regions of interests for SNR calculations throughout all bands hence it is assumed to contain a high SNR level. This is required, due to the fact that the noise may dominate the signal in the peripheral bands. Having the middle band as the reference band we apply K-means clustering method for segmentation of the region of interests (RIO) representing homogeneous regions. The cluster/region with a maximum number of pixels (i.e. largest connected component) which may correspond to a uniform background or skin region is selected for SNR calculation. The

largest region is declared as the mask marking the homogeneous region which is used to calculate the SNR for each spectral band. The SNR value is used to calculate the band-specific weight values.

2. In the second contribution, all bands are fused to a single band by multiplying each band with a weigh based on a band-specific SNR values. This method converts 3D hyperspectral face cubes into 2D face images which contain contributions from each spectral subbands through the method we developed. Generated 2D images go through PCA for dimensionality reduction. The feature vectors result from PCA are classified by using *k*-NN and CRC classifiers.

4.2 Signal to Noise Ratio (SNR) Estimation

In 3D hyperspectral image cubes, both spatial and spectral correlations are included [54]. Hyperspectral face image can be represented by a matrix of true unknown signal combined by additive noise. High dimensionality and low signal to noise ratio are the challenges which are posed by hyperspectral images. Low signal to noise ratio appears in certain spectral bands which are located near the blue wavelength.

Signal to Noise ratio (SNR) is a mathematical method for quality analysis of hyperspectral images. There is a problem to compute a metric in hyperspectral images since they do not have reference data to be compared with [49]. Linlin Xu [55] proposed the SNR calculation for a hyperspectral image, for a given band *k*. He proposed to calculate the SNR as follow.

$$SNR_k = 10log_{10}\frac{\sum_{ij}\left(I_{ij}^k\right)^2}{\sum_{ij}\left(I_{ij}^k - \mu^k\right)^2} \quad (4.1)$$

$$\widehat{SNR}_k = \frac{SNR_k}{\sum_{k=1}^{\beta} SNR_k} = \left[\widehat{SNR}_1, \widehat{SNR}_2, , \ldots . , \widehat{SNR}_\beta\right] \quad (4.2)$$

Where I_{ij}^k denotes the pixel at spatial location i, j in an area where pixels are homogeneous in the k^{th} band in spectral domain. The mean value of I_{ij}^k is defined by μ^k. The selection of homogeneous area is important since SNR estimation depends on this area. Normalized SNR for each band is given in Equation 4.2.

4.3 K-means Clustering Algorithm

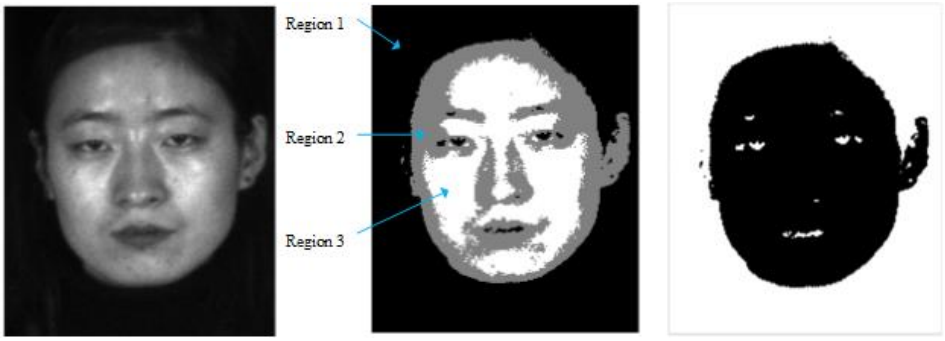
Cluster analysis (data clustering) discover the standard grouping of a set of patterns, points, or objects. Data clustering is an unsupervised learning or classification which means there is no predefined class or label. A collection of data in a cluster are Similar to one another and share common characteristics and dissimilar to the data in other clusters. Clustering algorithm group unlabeled data (object) which are considered to be similar based on common characters or features [50].

K-means clustering which is a type of unsupervised learning is one of the easy and well-known algorithms for grouping data/objects [51]. This algorithm based on feature similarity, will find groups in the data, and the number of groups (initial centroids) are represented by the variable K [53]. K-means algorithm assign each data point to the closest centroid based on features similarity iteratively. The position of centroid for each cluster is moved by the means of the data points assigned to center. This procedure repeats until no centroid is shifted in an iteration which means the minimum shift is below the threshold. The grouping is done when the distances between data and its cluster centroid is minimized. K-means clustering algorithm because of its simplicity and high speed has an advantage of processing on big data [52].

4.4 Proposed Method Using K-means Clustering to Estimate SNR

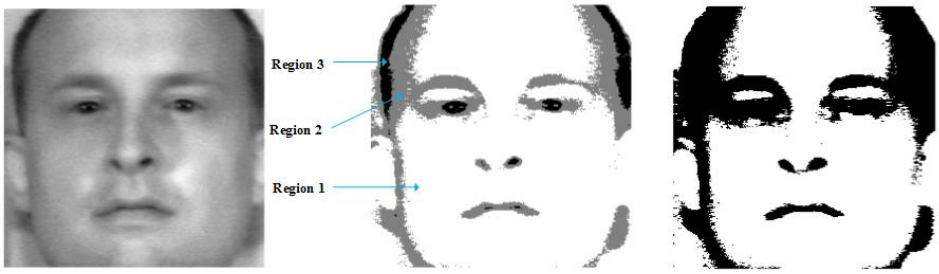
Calculating SNR in a region which is constant or with low frequency signal content which is regarded as homogeneous is important. In a homogeneous region, the additive noise dominates the signal since the signal is of low frequency nature. In this regard, homogeneous regions are of our region of interest for the calculation of SNR values for each band. Determining such regions require segmentation. In this chapter, K-means based clustering is used to segment the homogenous region of interests.

In this approach, we assume that the middle band in hyperspectral face image has higher SNR values since they are far away from blue and infrared bands which have lower SNR [49]. Hence the middle band is used for the segmentation of the region of interest for SNR calculation. K-means clustering is applied with, K=3 and the cluster with the largest number of pixels are chosen the homogeneous region of interest. A mask with zero and one value is created in this band where one is given to each pixel in the homogeneous region and zero to all pixels in other clusters. The mask for a subject in PolyU, CMU and UWA databases are shown in Fig. 4.1(c), Fig. 4.2(c) and Fig. 4.3(c) respectively. Multiplying each spatial coordinate (i, j) in each band with this mask help to segment the homogeneous region in each band. The illustration of each band after applying mask for a subject in PolyU, CMU and UWA databases are shown in Fig. 4.4, Fig.4.5 and Fig. 4.6 respectively. The SNR value is calculated for each band of hyperspectral face image using Eq. 4.1 and normalized by Eq. 4.2.



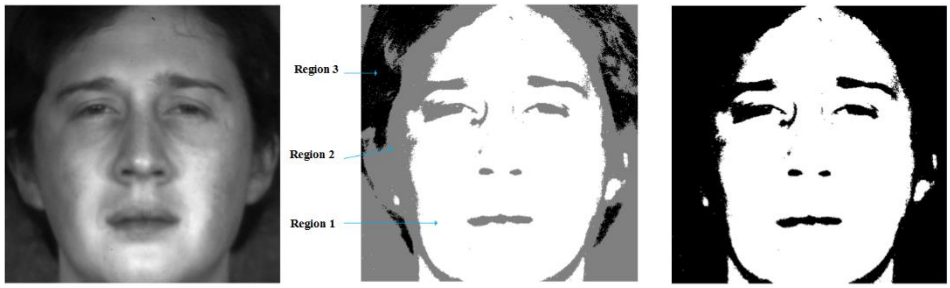
(a) (b) (c)

Figure 4.1: (a) the middle band, (b) three segmented region, (c) created mask for a subject in hyperspectral face cube from the Poly U-HSFD.



(a) (b) (c)

Figure 4.2: (a) the middle band, (b) three segmented region, (c) created mask for a subject in hyperspectral face cube from the CMU-HSFD.



(a) (b) (c)

Figure 4.3: (a) the middle band, (b) three segmented region, (c) created mask for a subject in hyperspectral face cube from the UWA-HSFD.

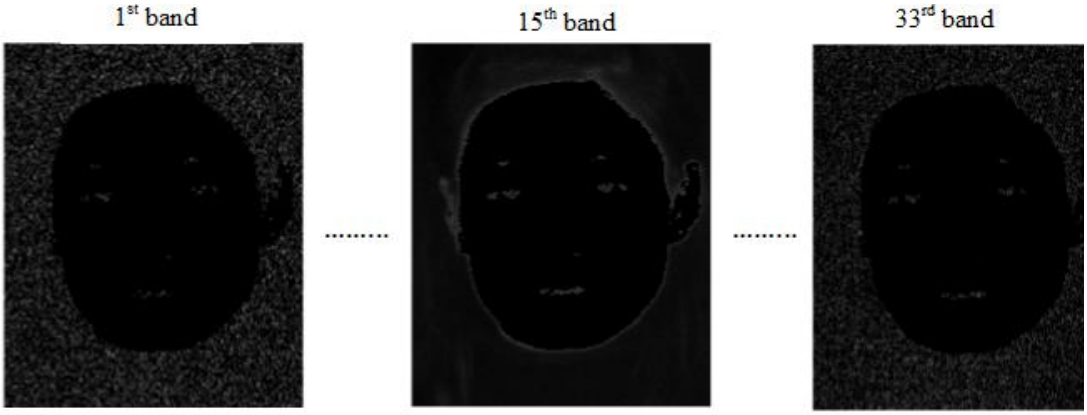


Figure 4.4: The illustration of each band after applying mask for a subject in PolyU-HSFD.

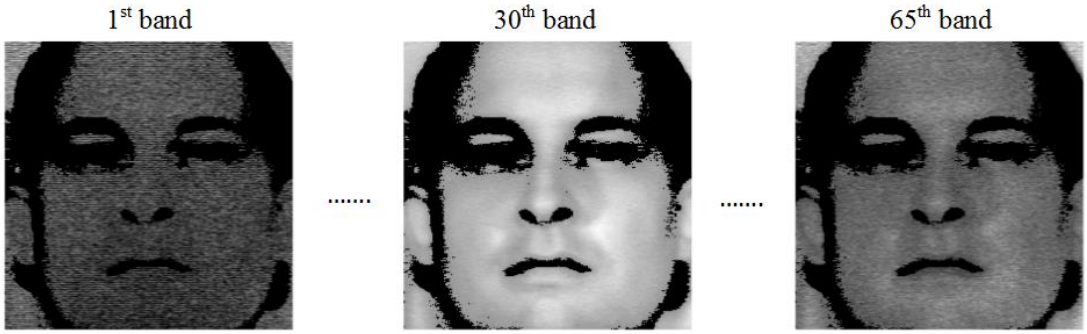


Figure 4.5: The illustration of each band after applying mask for a subject in CMU-HSFD.

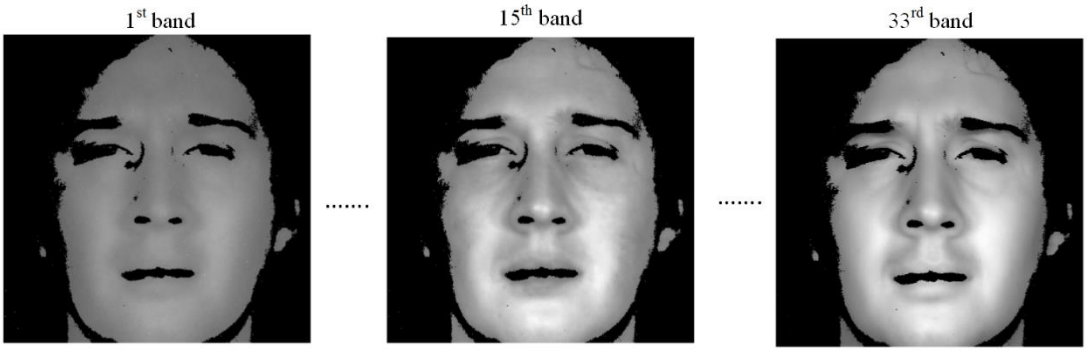


Figure 4.6: The illustration of each band after applying mask for a subject in UWA-HSFD.

The normalized $S\widehat{N}R$ corresponding to each band is used as a band-specific weight of each spatial coordinate (i, j) in the respective band. According to $S\widehat{N}R$ value of each band, the bands with more information have more weight. All pixels with spatial

coordinate (i, j) along spectral direction β are added and results 2D face image instead of 3D face cube. Following the fusion of the 3D face cubes, 2D face images are cropped according to the eye coordinates and resized to 64×64 [44]. Visualization of the fusion process of the 2D face image for a subject in PolyU, CMU and UWA databases are shown in Fig. 4.7, Fig. 4.8 and Fig. 4.9 respectively. Each cropped 2D image is applied to PCA for dimensionality reduction, then they are applied to *k*-NN and CRC classifiers. Fig. 4.10, Fig. 4.11 and Fig. 4.12 illustrates the $\widehat{SNR}$ value of each band of a subject from PolyU, CMU and UWA databases.

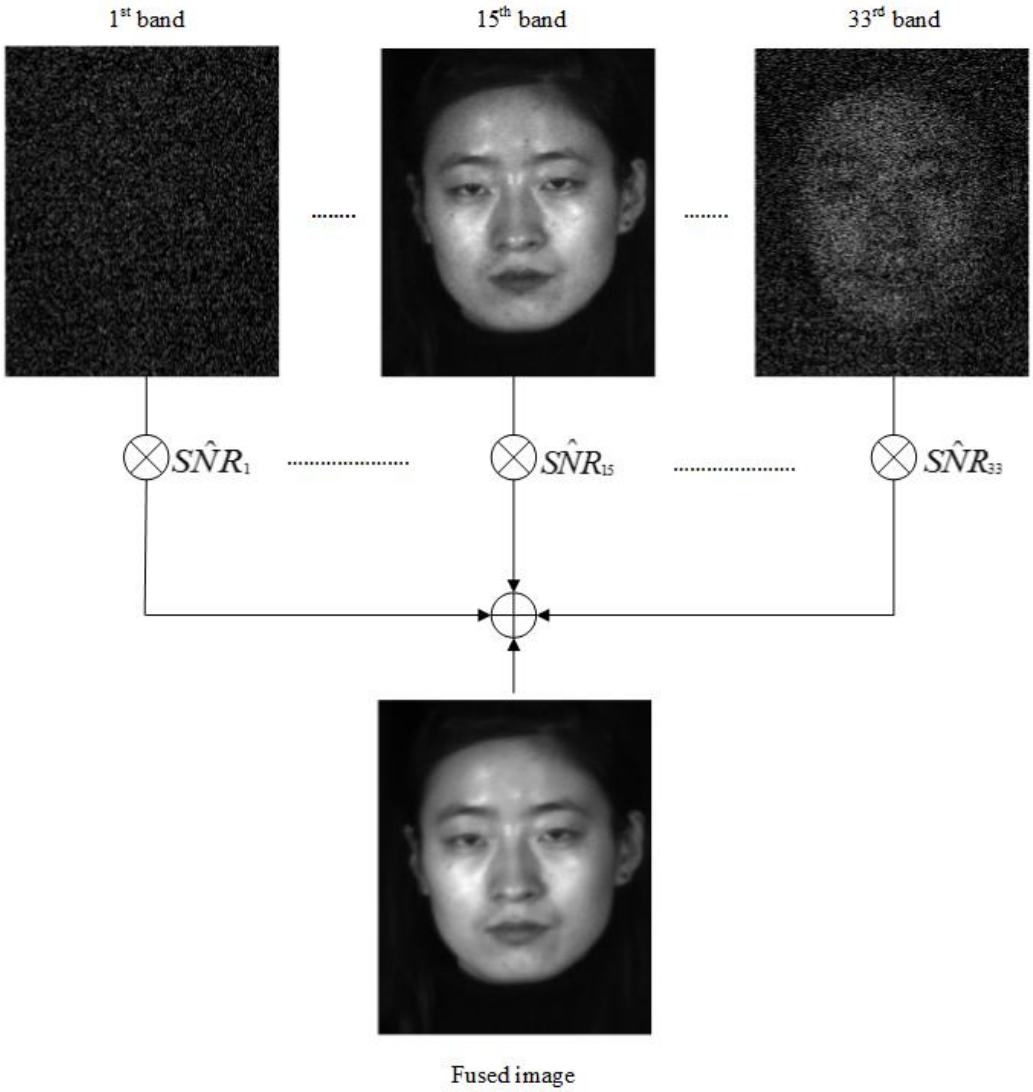


Figure 4.7: Fusion process of the 2D face image for a subject from PolyU-HSFD.

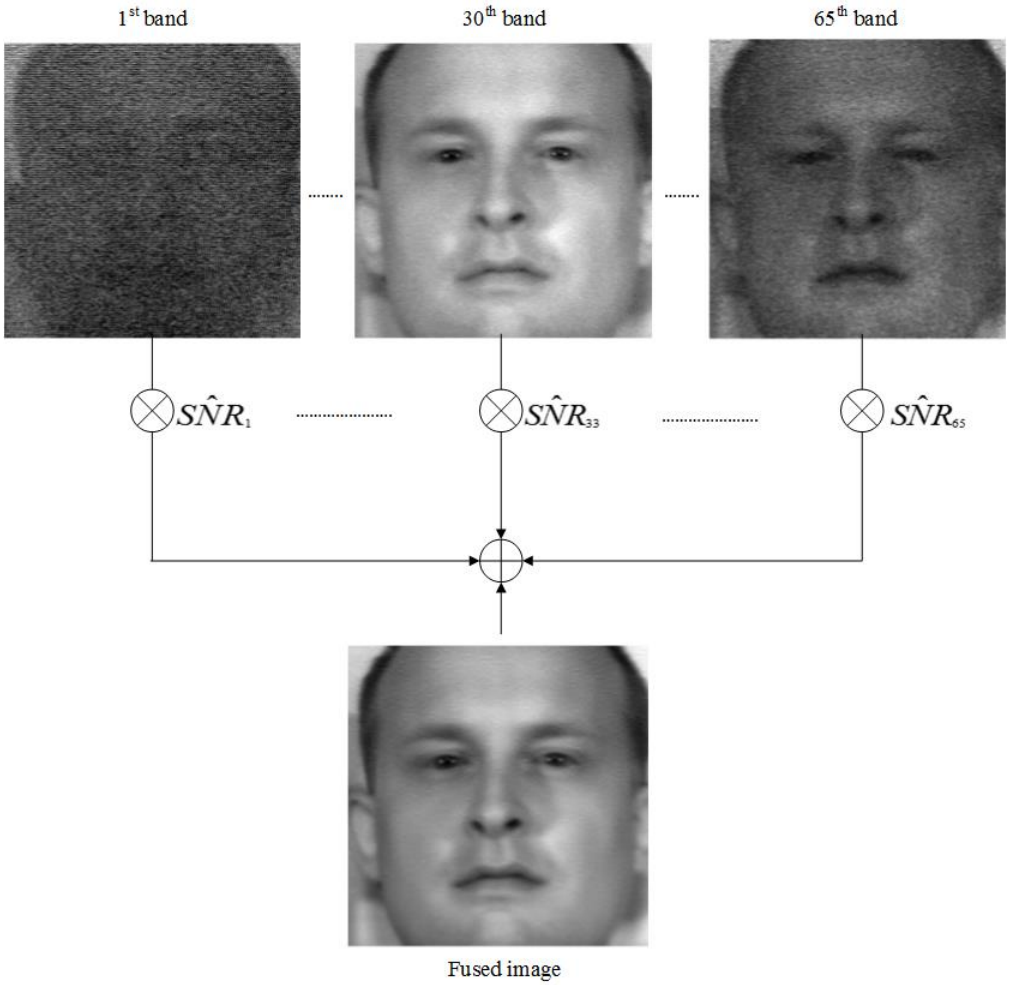


Figure 4.8: Fusion process of the 2D face image for a subject from CMU-HSFD.

Figure 4.9: Fusion process of the 2D face image for a subject from UWA-HSFD.

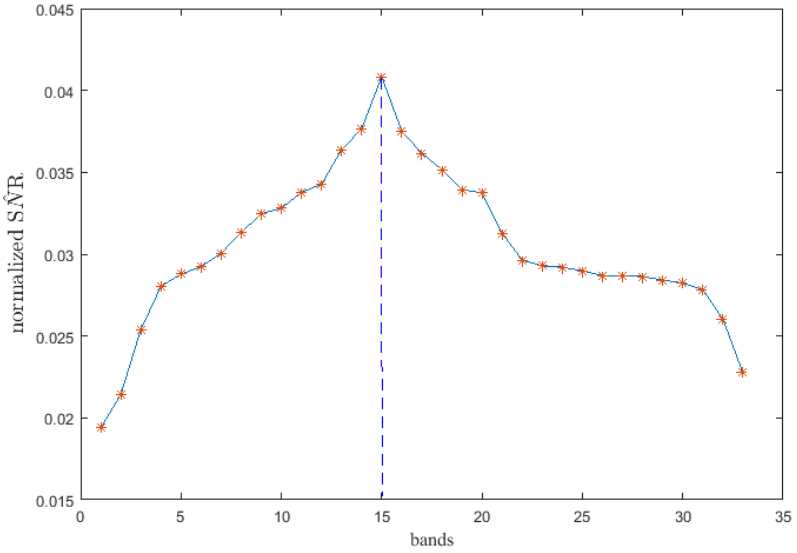


Figure 4.10: $\widehat{SNR}$ value for each band of a hyperspectral face cube from the Poly U-HSFD.

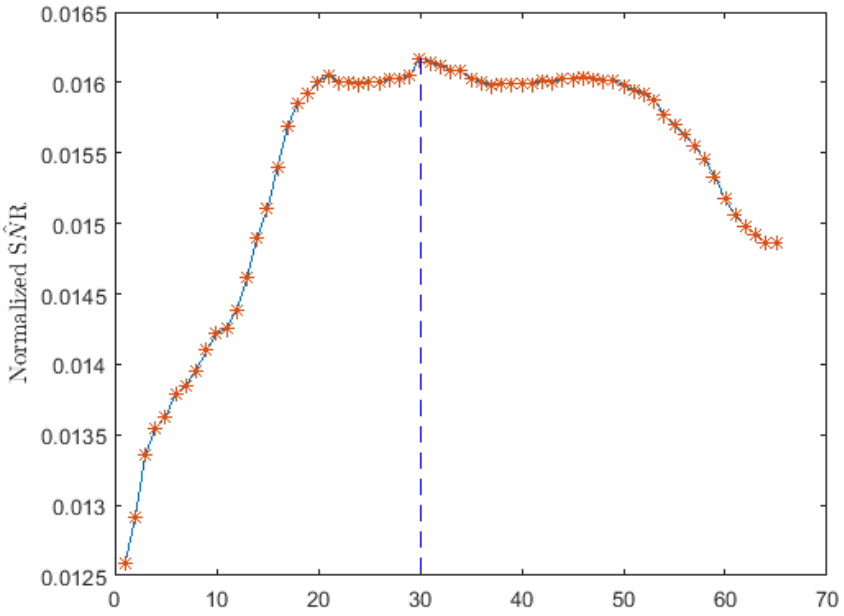


Figure 4.11: $\widehat{SNR}$ value for each band of a hyperspectral face cube from the CMU-HSFD.

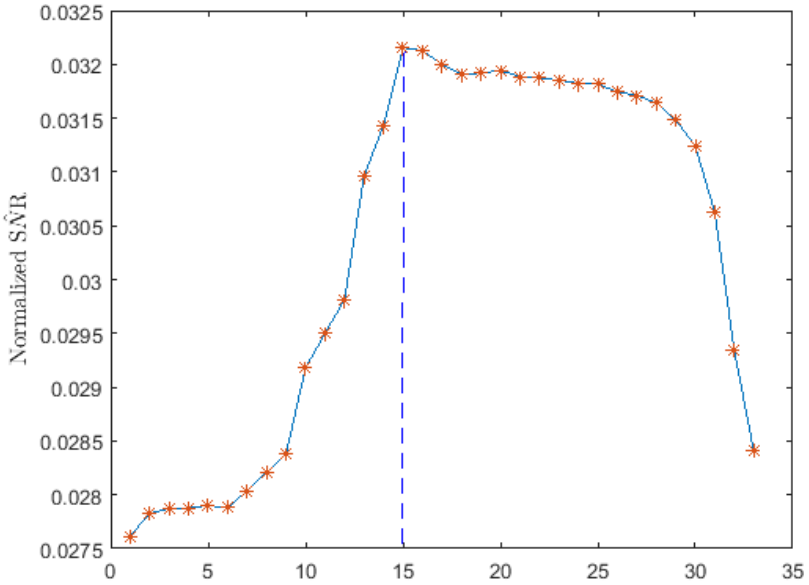


Figure 4.12: $\widehat{SNR}$ value for each band of a hyperspectral face cube from the UWA-HSFD.

4.5 Experimental Results and Comparison

The experiments are performed on three hyperspectral face databases (HSFD), PolyU-HSFD [4], CMU-HSFD [5] and UWA-HSFD [6] which are explained in details in section 3.2 to validate the performance of the proposed methods.

4.5.1 Testing Set Distribution

The experimental methodology suggested by [11, 13] is adopted for PolyU-HSFD, where first 25 subjects of frontal face images with all 33 spectral bands are used in the experiment. For each subject, 2 cubes are selected randomly for training. The rest of the 63 cubes are used for testing. In CMU-HSFD, as suggested by [11], we selected the session for each subject with all lights on which contain 1-5 cubes. One cube is randomly chosen for training set and the rest cubes as testing set for each subject. Following the experimental methodology of [11], in UWA, for each subject one cube is randomly selected for training and the rest 50 cubes as testing set.

4.5.2 Experimental Results Comparison

The proposed methods are performed ten times to create training and testing sets randomly to generate the average accuracy and standard deviation (STD). The accuracy of correct classification is defined as the correctly classified hyperspectral faces over the total number of hyperspectral faces in the test set. For classification *k*-NN and CRC classifiers are adopted which are explained in details in section 3.6.1 and 3.6.2.

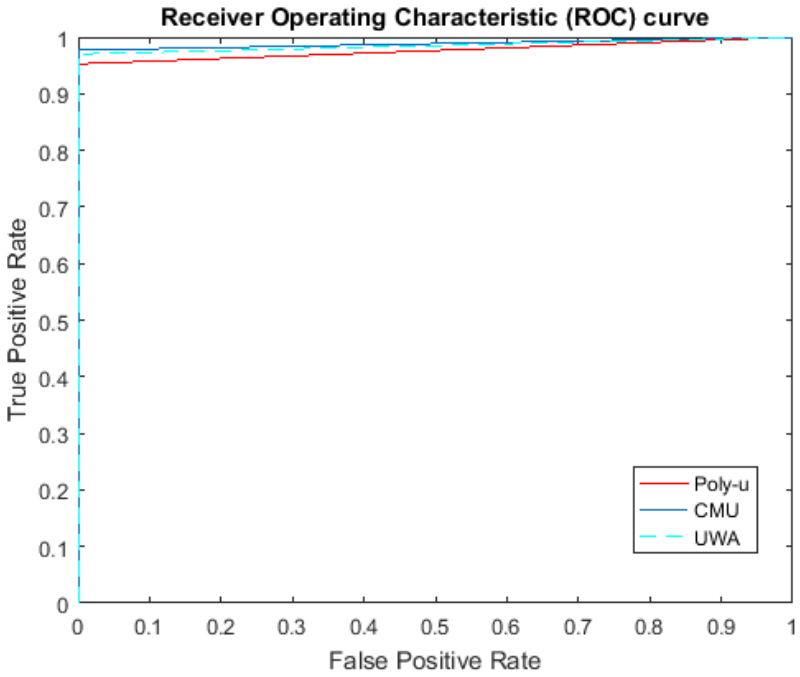


Figure 4.13: ROC curve of the proposed SNR fusion method, for all three standard databases by adopting *k*-NN classifier.

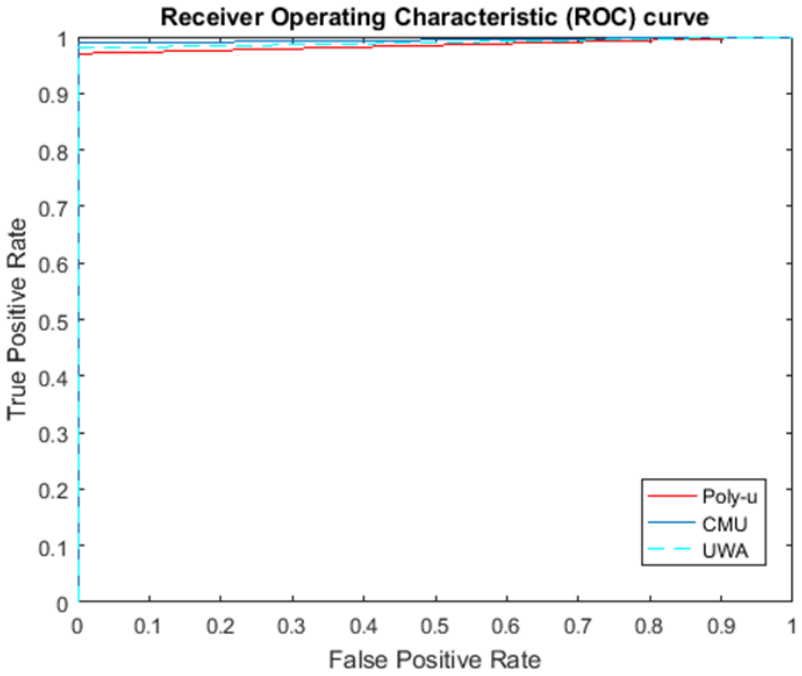


Figure 4.14: ROC curve of the proposed SNR fusion method, for all three standard databases by adopting CRC classifier.

The Receiver Operating Characteristic (ROC) curves for SNR based fusion method for all three databases for *k*-NN and CRC classifiers are illustrated in Fig. 5.13 and Fig. 5.14 respectively. The ROC curves show the True Positive Rate (TPR) against False Positive Rate (FPR). Area Under the Curve (AUC) indicated the classifier hypersurface. However, by extending two-class ROC to multiclass the visualization can be possible [42]. This approach has two forms; *one versus all* and *one-versus-one*. In *one-versus-all* form, the classification performance is evaluated for class '*one'* and all the remaining classes are considered into a single class *'not on*e'. Therefore, details of specific misclassification errors of the class *'not one'* are not available [15]. In this paper, we have adopted *one-versus-all* approach. Figure 4.14 shows that CRC classifier comparing to Figure 4.13 *k*-NN classifier for all three standard databases has higher recognition accuracy since its AUC is closer to one.

Comparisons of proposed method with several existing methods for all three databases are shown in Table 4.1. The accuracy of proposed SNR based band fusion method by adopting *k*-NN and CRC classifiers reach to 95.71%±2.3 (97.09%±1.32), 97.7%±1.91 (98.93%±0.65) and 97.02%±1.63 (98.52%±1.24) for PolyU-HSFD, CMU-HSFD and UWA-HSFD databases respectively. The proposed method outperforms alternative methods of the state-of-the-art.

Table 4.1: Recognition accuracy for proposed SNR band fusion method.

Databases	Gabor wavelet [15,26]	Log-polar FFT2 [15]	3D LDP [10]	Band fusion +PLS [26]	Feature extraction and CRC [16]	3D-GE	Proposed DWT based band fusion method (Max Pooling) (CRC)	Proposed SNR based band fusion method (*k*-NN)	Proposed SNR based band fusion method (CRC)
PolyU	91.3%±2.1	94.6%±2.5	95.3%±1.6	95.2%±1.6	96.4%±2.3	96.66%±1.2	97.07%±1.78	95.71%±2.3	**97.09%±1.32**
CMU	91.6%±2.9	95.6%±1.7	94.8%±2.6	**99.1%±0.6**	98.0%±0.7	98.61%±1.3	98.88%±0.42	97.7%±1.91	98.93%±0.65
UWA	91.5%±3.07	**-**	**-**	98.2%±1.2	**-**	98.28%±1.05	98.37%±1.13	97.02%±1.63	**98.52%±1.24**

In Table 4.1, it can be seen that the results obtained from the proposed SNR based band fusion method using CRC classifier has higher recogniion rate comparing to same method using *k*-NN classifier. CRC has an advantage of classifying each testing image to the class with minimal regularized reconstruction error instead of classifying image by a majority vote that is given by the k neighbors of it in *k*-NN classifier. Proposed SNR based band fusion method comparing to our other two contributions has higher recognition accuracy since SNR value is used to calculate the band-specific weight values, higher SNR value means more information.

In this contribution SNR value is used to calculate the band-specific weight which has a computational complexity of $\mathcal{O}(N^2)$.where $N$is an input's dimension. In [13], they employed 3D-Gabor wavelet to extract features of hyperspectral face images which has the same computational complexity of $\mathcal{O}(N^3)$. In [15-16], features are extracted by 2D Fast Fourier Transform (FFT) from the log-polar images hence the computational complexity of FFT is $\mathcal{O}(NlogN)$. In [26], the first and second order statistics of each cubelets (section 2.4.2) are computed by the mean vector and covariance of 2D matrix which covariance matrix computation has a computational complexity of $\mathcal{O}(N^2)$. Computational complexity of proposed method is same as [26], better than [13] and higher than [15-16].

4.6 Conclusion

In this chapter, we propose a novel method to fuse spectral bands in hyperspectral face image cubes to improve hyperspectral face recognition. The proposed method is able to overcome some of the challenges which are posed by hyperspectral face images such as high dimensionality and low signal to noise ratio. Proposed methods overcome these problems without losing information from spectral bands. In this method, each

3D face cube is fused along different spectral bands into a 2D image by using band-specific SNR based weighting. Weights based on the estimated $\widehat{SNR}$ values scale the pixels intensities in each band, hence each pixel along spectral axis is fused to a single pixel. Each 2D face image go through PCA for dimensionality reduction and feature vector creation before *k*-NN and CRC based classifications. The experimental results using the SNR based weighting method show that, the performance in terms of classification accuracy outperforms several state-of-the-art methods using standard hyperspectral face databases, PolyU- HSFD, CMU-HSFD and UWA-HSFD.

Chapter 5

DWT BASED FUSION ALONG SPECTRAL AXIS TO IMPROVE HYPERSPECTRAL FACE RECOGNITION

5.1 Introduction

As mentioned in previous chapters, facial hyperspectral images samples are captured at multiple narrow bands within the neighborhood of visible light in the electromagnetic spectrum, which reveals significant information that is not evident in RGB/grey images. This work is specially intended to benefit from the available rich 3D-cube information to create more competitive results than 2D methods. RGB images are special version of hyperspectral images where the visible band is only quantized into three subbands (Red-Green-Blue). In other words, hyperspectral images not only include finer spectral details of the image in the visible band but also extends into near-infrared or ultraviolet bands importing more information than RGB images. Therefore, from the information theory perspective hyperspectral images in the form of 3D-cubes definitely contain more information.

In this chapter, we propose a new hyperspectral face recognition method using spectral band fusion. The proposed method extracts a single 2D image from a given hyperspectral image cube containing multiple spectral bands. The method effectively fuses spectral information by applying discrete wavelet transform (DWT) to each pixel along the spectral axis on each hyperspectral image cube. The DWT is applied consecutively to each vector of pixels extracted from the low-frequency component

from the previous decomposition. This process is iteratively repeated until the spectral vector for each pixel is decimated to a single pixel. Once all the pixels go through the same operation, the 3D input spectral image cube is transformed into a 2D output image. This operation can be regarded as a fusion since it transforms a hyperspectral image cube into a single 2D image. Generated 2D images go through Principle Component Analysis (PCA) for dimensionality reduction. The feature vectors result from PCA are classified by using *k*-NN and CRC classifiers. The results are compared with alternative methods in the literature. The performance of the proposed method in terms of accuracy outperforms existing methods using PolyU [4], CMU [5] and UWA [6] Hyperspectral Face Databases.

This chapter includes two contributions which are listed below:

1. The first contribution involves fusing spatio-spectral information by applying discrete wavelet transform (DWT) to each pixel along the spectral axis on each hyperspectral image cube. The DWT is applied consecutively to each spectral vector of pixels extracted from the low-frequency component from the previous decomposition. This process is iteratively repeated until the spectral vector for each pixel is decimated to a single-pixel transforming the 3D input spectral image cube into a 2D output image.
2. The second contribution is about generating the feature vectors by applying principal component analysis (PCA) to transformed 2D images.

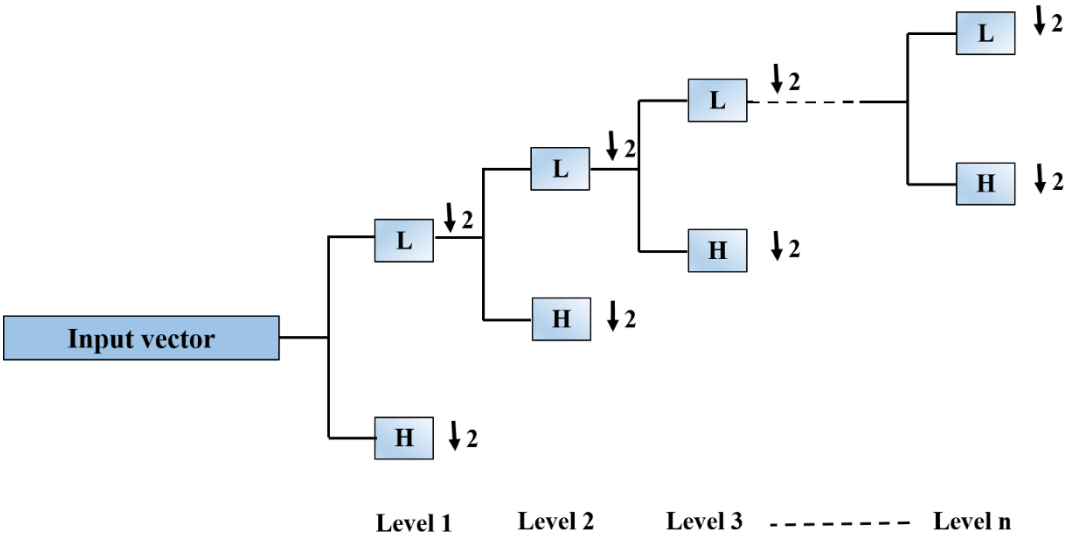


Figure 5.1: *n* level DWT procedure through lowpass filter bank.

5.2 One Dimensional Discrete Wavelet Transform (DWT)

Wavelet Transform provides a possibility of analyzing signals in time (spatial) and frequency domains simultaneously. Hence every frequency component is not resolved equally [43]. To analyze signal at different resolutions and scales, filtering and sampling operations are used respectively. DWT employs scaling and wavelet functions to analyze the signal at different frequency bands with different resolutions. Signals can be decomposed into different frequency bands by low and high pass filtering of the time domain signal. The original signal is first passed through a highpass filter H and a lowpass filter L. After filtering, the signal has a frequency of $\pi/2$ instead of π hence half of the samples according to the Nyquist's rule can be removed [43].

DWT operates on signals which can be considered as a vector of real values, with the size of length 2^n, $n \in \{2, 3\ldots\}$. The resulting vector has the same length as the original signal. This vector has two filtered parts, the first filtered part which has the length $2^n/2$ includes the coefficients resulting from low-pass filtering (LPF) of the original

signal and the second filtered part which has the length $2^n/2$ includes the coefficients resulting from high-pass filtering (HPF) the original signal. The DWT is regarded as a multiresolution analysis which can analyze resolutions and scales in 1D discrete-time signal. For further decomposition, the procedure can be repeated *n* level. Fig. 5.1 shows *n* level DWT filter bank showing the consecutive process through lowpass filtering.

The first and simplest orthonormal wavelet basis is Haar wavelet. The Haar transform calculates pair wise averages and differences since it uses just two scaling and wavelet function coefficients. As in this transform there are no overlapping windows, only changes between adjacent pixels pairs can be reflected.

The most popular wavelet family which is used for texture feature analysis is Daubechies wavelet family. Unlike Haar wavelet, Daubechies wavelet uses overlapping windows hence high frequency coefficient spectrum reflects all high frequency changes. Therefore, Daubechies wavelets are useful in noise removal [56].

5.3 Principle Component Analysis

The aim of principle component analysis (PCA) is to explain the variation in the set of the training set by a few variables. The most famous method based on PCA is an eigenface method [45]. In this method instead of analyzing the whole face image, analyzing the difference between individual face will be enough.

Let the face image be the size of $M \times M$. To calculate eigenfaces, it is necessary to find a basis which is meaningful enough to re-express an ensemble face image [46]. Let training set be $\gamma = [\gamma_1, \gamma_2 \ldots, \gamma_N]$. To produce a training set whose mean is zero,

each face differs from a mean face of training set [47]. Mean face and the vector of each face differs from the mean face are defined respectively as,

$$\boldsymbol{\mu} = \frac{1}{N}\sum_{i=1}^{N} \gamma_i \tag{5.1}$$

$$\boldsymbol{\psi}_i = \boldsymbol{\gamma}_i - \boldsymbol{\mu} \tag{5.2}$$

The aim of analyzing covariance is to find out how much the dimensions vary from the mean with respect to each other. The sign of covariance is more important than its exact value. Positive value indicates that both dimensions increase together. Negative value means as one dimension increases, the other decreases and zero covariance shows that the two dimensions are independent of each other [47]. After calculating all the possible covariance values between all the different dimensions, they are placed in a matrix which is in general presented by equation (5.3) [13]. In this case, it is better defined it as an equation (5.4)

$$C^{n\times n} = (c_{i,j}, c_{i,j} = cov(Dim_i, Dim_j)) \tag{5.3}$$

$$C = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{\psi}_i \boldsymbol{\psi}_i^T = AA^T \tag{5.4}$$

where $A = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \dots, \boldsymbol{\psi}_N]$.

Next step is finding the eigenvectors and eigenvalues of the covariance matrix which can differentiate the face images. Hence the eigenvectors and eigenvalues of the covariance matrix are the principal component of the training set. In general, the number of eigenfaces is equal to the number of face images in the training set. In large databases it means a lot of processing. To overcome this problem, it is possible to represent face images by using the best eigenfaces which are the ones with the largest eigenvalue. The highest eigenvalue with respective eigenvector which has the best

description of the data distribution is going to be used. Hence the k^{th} eigenvector u_k with maximum λ_k is chosen [48].

$$\lambda_k = \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{u}_k^T\boldsymbol{\psi}_i)^2 \tag{5.5}$$

M^2 eigenvectors and eigenvalues are determined by the covariance matrix of dimensions $M^2 \times M^2$. It can be considered $N-1$ meaningful eigenvectors instead of M^2 significant eigenvectors if the dimension of image space is greater than the number of data points $N \ll M^2$ [45]. The remaining eigenvectors are associated to the eigenvalues which have the value close or equal to zero. N eigenvectors (V_l) from the ($N \times N$) matrix $L = \psi\psi^T$ where $L_{mn} = \psi_m^T\psi_n$ can be constructed by this analysis. The N training set of face images in the form of eigenfaces is determined by linear combinations these vectors [48],

$$u_l = \sum_{k=1}^{N}\boldsymbol{V}_{lk}\boldsymbol{\psi}_k \qquad\qquad l = 1, \dots, N \tag{5.6}$$

A face image (γ) is transformed into its eigenfaces components which means it is projected into face space by the following operation:

$$w_p = \boldsymbol{u}_p^T(\boldsymbol{\gamma} - \boldsymbol{\mu}) \tag{5.7}$$

5.4 Proposed Method for Fusion of Hyperspectral Face Cube along Spectral Axis by DWT

Nowadays hyperspectral face recognition is a new topic in biometrics applications. Facial hyperspectral images have more information comparing to traditional 2D images according to different response of face in each spectrum. In this chapter, we propose a new approach in which, for each hyperspectral face cube, 1D-DWT is applied to vectors that are generated by each pixel along spectral axis. By considering the low frequency part after DWT downsampling, in each level the number of pixels

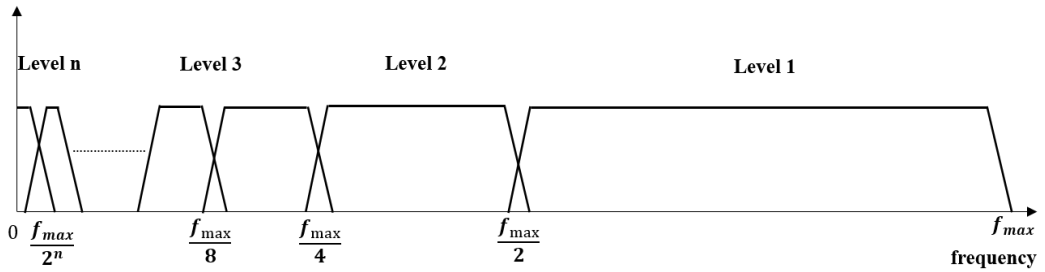


Figure 5.2: Frequency progress for *n*-level decomposition.

which characterize the entire image decrease to half. Consecutive analysis of the low frequency part of the signal in the spectral axis of a hyperspectral image cube spans the half of previous level. Fig. 5.2 illustrate *n* level frequency spanning in DWT decomposition. By applying the same decomposition consecutively to lowpass component of each level, the last level results a single coefficient for each spectral vector for every spatial coordinate generating a 2D image instead of 3D cube. The number of levels is calculated by $n = log\, N$ where *N* is the length of spectral bands in the hyperspectral face cube.

In addition, we propose to adopt Maximum Pooling (Max Pooling) method, which is used in Convolutional Neural Networks [64] providing an approach to down sample feature maps by summarizing with the most activated presence of a feature [65]. This down sampling approach is shown in Fig. 5.3.

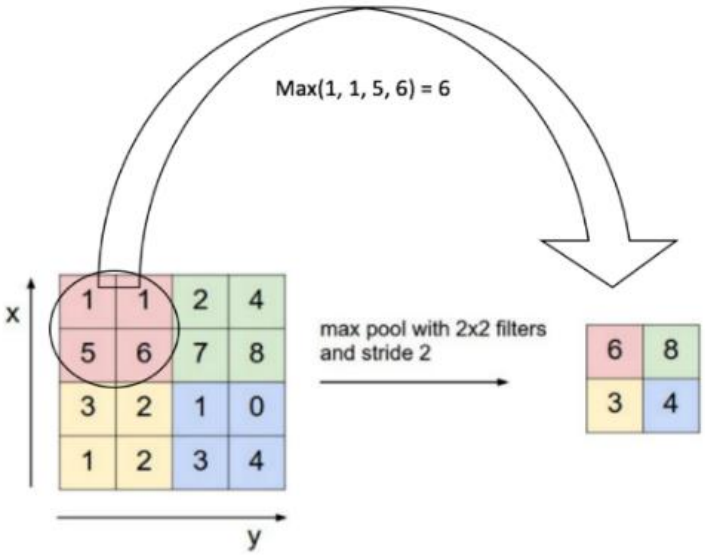


Figure 5.3: Max Pooling operation on feature map (2×2 window).

After applying the approach, where *n* level DWT is applied to each vector that is generated by each pixel along spectral axis up to the level of decomposition that decimated the vector length to 4, we have adopted Max Pooling method to the generated 4 wavelet coefficients. This 1×4 vector can be considered similar to a matrix of 2×2 and the maximum valued pixels is chosen as the most activated presence of a feature. Generated final coefficient is a summarized version of the features detected in the vector resulting a 2D image instead of 3D cube for each spectral vector for every spatial coordinate. These 2D images are cropped according to the eye coordinates and resized to 64×64 [44]. Each cropped 2D image is applied to PCA for dimensionality reduction and *k*-NN and CRC use for classification. The proposed method is illustrated in Fig. 5.4.

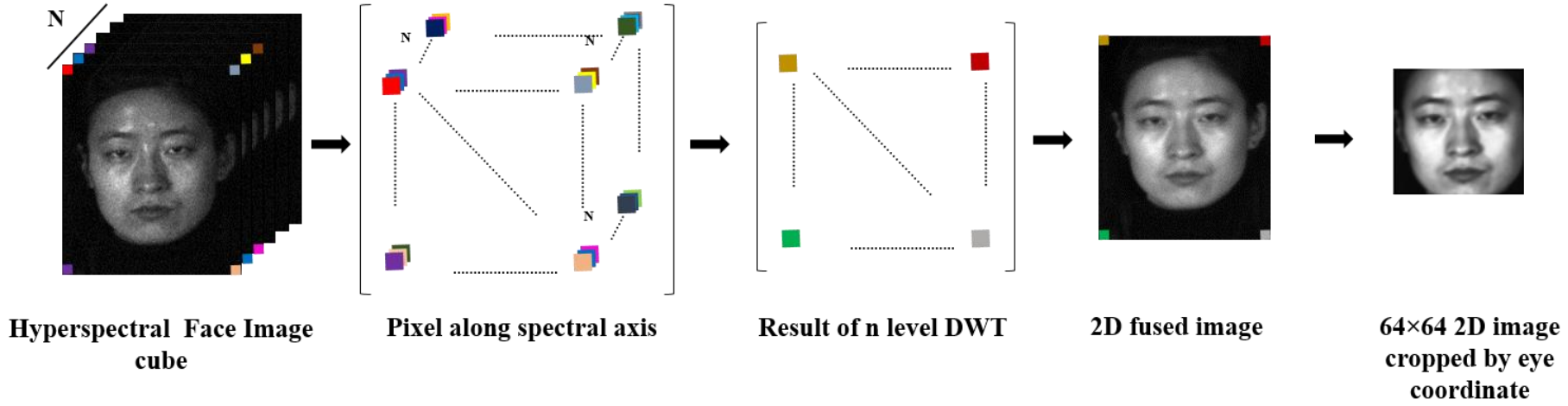


Figure 5.4: Proposed DWT based fusion along spectral axis.

The proposes DWT based fusion which transforms a given 3D Hyperspectral image cube into a 2D image can be regarded as a smoothing operation in the spectral axis. Having a signal low-pass filtered *n* times using DWT based filter bank effectively removes the inherent noise added through the acquisition process of the hyperspectral imaging systems. Noise removal by DWT along spectral axis is the key for enhanced fusion framework proposed in this chapter. The operation is repeated *n* ($n = log\, N$) times to fuse *N* coefficients of each spectral vector into a single coefficient for each pixel spatial coordinate. The method can be formulated as,

$$\delta_{low}(x, y) = \sum_{z=1}^{N} C(x, y, z) h(2n - z) \tag{5.8}$$

$$\delta_{high}(x, y) = \sum_{z=1}^{N} C(x, y, z) g(2n - z) \tag{5.9}$$

Where $\delta_{low}(x, y)$ and $\delta_{high}(x, y)$ are the outputs of decomposed hyperspectral face cube after passing through highpass and lowpass filters for one level respectively. $C(x, y, z)$ is the hyperspectral cube with *x* and *y* as spatial and *z* as a spectral dimensions. $h(z)$ and $g(z)$ are lowpass and highpass filters kernel respectively. As illustrated in Fig. 4.1, only low pass filter is used in a consecutive manner downsampling the signal by 2 which halves the signal (spectral vector) size after each iteration. This operation is repeated *n* (*n*=*log N*) times until a single sample is obtained.

Alternative smoothing techniques can be proposed to transform the 3D Hyperspectral image cube into a 2D image. Simple averaging operation,$\alpha(x, y)$, for each pixel over the coefficients along the spectral axis would also generate a 2D image as follows:

$$\alpha(x, y) = \frac{1}{N} \sum_{z=1}^{N} C(x, y, z) \tag{5.10}$$

Similarly, energy along the spectral axis is another form of smoothing which can be used to fuse the 3D image cubes into 2D image. The energy value, $\varepsilon(x, y)$, for each pixel can be calculated by,

$$\varepsilon(x, y) = \sum_{z=1}^{N} C(x, y, z)^2 \tag{5.11}$$

(a) (b) (c)

Figure 5.5: (a) average/db1, (b) energy, (c) proposed method for a subject from PolyU-HSFD.

(a) (b) (c)

Figure 5.6: (a) average/db1, (b) energy, (c) proposed method for a subject from CMU-HSFD.

(a) (b) (c)

Figure 5.7: (a) average/db1, (b) energy, (c) proposed method for a subject from UWA-HSFD.

Visualizations of the proposed DWT based fusion, fusion by averaging and fusion by energy along spectral axis for PolyU-HSFD, CMU-HSFD and UWA-HSFD are shown in Fig. 5.5, Fig. 5.6 and Fig. 54.7 respectively. In part (a)'s all images are blurred which are the average of an image along spectral axis, part (b)'s images are dark which are energy of an image (globally best value) along spectral axis and in part (c)'s all images are sharper since Max Pooling keeps the maximum valued pixels which is called as most activated presence of a feature.

5.5 Experimental Results

In order to validate the performance of the proposed method, we perform experiments on three hyperspectral face databases (HSFD): PolyU-HSFD [4], CMU-HSFD [5] and UWA-HSFD [6] which are explained in details in section 3.2.

5.5.1 Testing Set Distribution

The experimental methodology suggested by [11, 13] is adopted for PolyU-HSFD, where the first 25 subjects of frontal face images with all 33 spectral bands are used in the experiment. For each subject, 2 cubes are selected randomly for training. The rest of the 63 cubes are used for testing. In CMU-HSFD, as suggested by [11], we selected the session for each subject with all lights on which contain 1-5 cubes. One cube is randomly chosen for training set and the rest cubes as testing set for each subject.

Following the experimental methodology of [11], in UWA, for each subject one cube is randomly selected for training and the rest 50 cubes as testing set.

5.5.2 Experimental Results and Comparison

The proposed method is performed ten times to generate the average accuracy and standard deviation (STD) by randomly creating training and testing datasets. The accuracy of correct classification is defined as the correctly classified hyperspectral faces over the total number of hyperspectral faces in the test set. For classification *k*-NN and CRC classifiers are adopted which are explained in details in section 3.6.1 and 3.6.2.

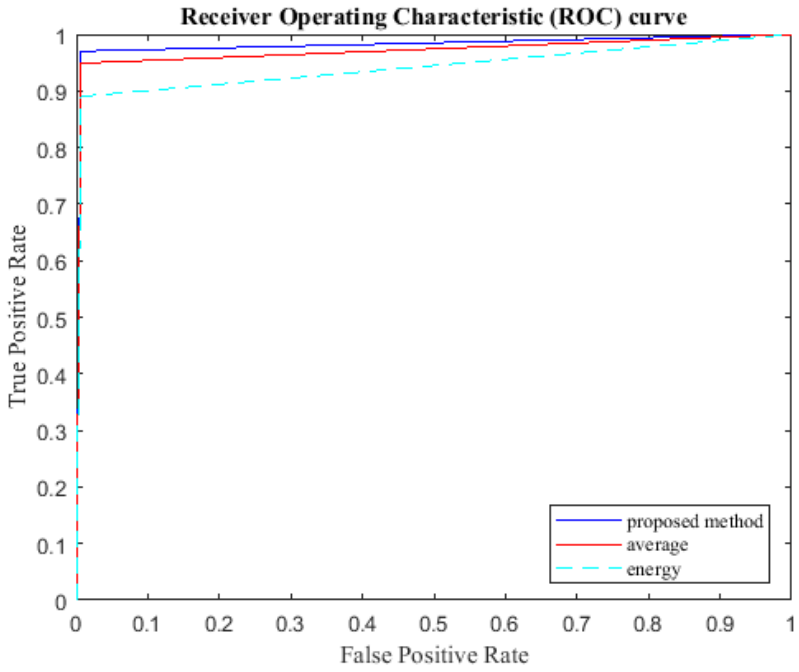


Figure 5.8: ROC curve of the proposed band fusion method, fused by energy and average for a subject from the PolyU-HSFD by adopting CRC classifier.

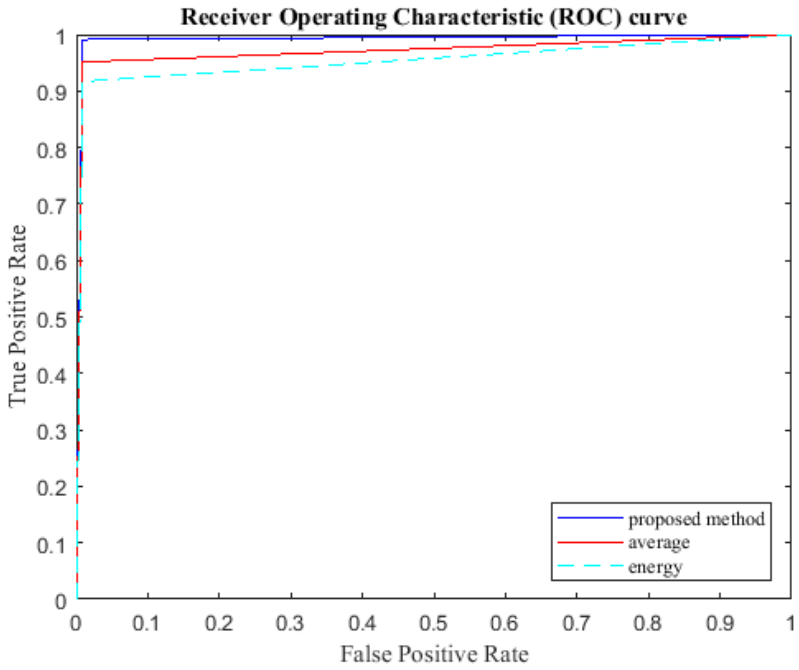


Figure 5.9: ROC curve of the proposed band fusion method, fused by energy and average for a subject from the CMU-HSFD by adopting CRC classifier.

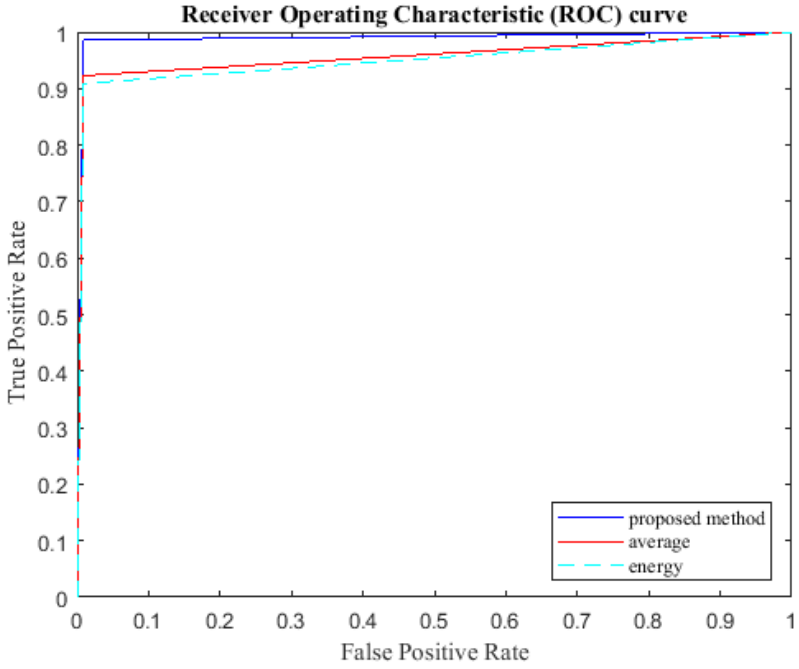


Figure 5.10: ROC curve of the proposed band fusion method, fused by energy and average for a subject from the UWA-HSFD by adopting CRC classifier.

The Receiver Operating Characteristic (ROC) curves for fused by energy, fused by average and fused by DWT methods for all three databases are illustrated in Fig. 5.8, Fig. 5.9 and Fig. 5.10 respectively. The ROC curves show the True Positive Rate (TPR) against False Positive Rate (FPR) which are explained in section 3.7.2. Area Under the Curve (AUC) indicated the classifier performance, where larger AUC implies the higher performance. The visualization of the ROC curve in multiclass problems is impossible. However, by extending two-class ROC to multiclass the visualization can be possible which are explained in section 3.7.2 [42]. In this chapter, we have adopted *one-versus-all* approach.

Comparisons of proposed method with several existing methods for all three databases are shown in Table 5.1. The accuracy of the proposed method reaches to 97.07%±1.78, 98.88%±0.42 and 98.37%±1.13 for PolyU-HSFD, CMU-HSFD and UWA-HSFD databases by adopting CRC classifiers respectively. The proposed method outperforms the work by Uzair *et al.* [26], which also use the concept of fusion. Additionally, the proposed method outperforms the rest five of the state-of-the-art methods tested in all three standard hyperspectral databases. The proposed method improves the recognition rates due to the repeated low-pass filtering along the spectral axis, by using DWT based filter bank, which removes the inherent noise added through the acquisition process of the hyperspectral imaging systems. Noise removal by DWT along spectral axis and fusing pixels along spectral axis by applying Max Pooling to the vector of 4 wavelet coefficients and choosing the most activated presence of feature are a key factor for the proposed fusion framework introduced in this chapter

Table 5.1: Recognition accuracy for proposed DWT based band fusion method.

Databases	**Gabor wavelet [15,26]**	**Log-polar FFT2 [15]**	**3D LDP [10]**	**Band fusion +PLS [26]**	**Feature extraction and CRC [16]**	**Proposed 3D-GE**	**Energy fused (*k*-NN/CRC)**	**Average fused/db1 (*k*-NN/CRC)**	**Proposed band fusion method (Max Pooling) (*k*-NN/CRC)**
PolyU	91.3%±2.1	94.6%±2.5	95.3%±1.6	95.2%±1.6	96.4%±2.3	96.66%±1.2	87.76%±2.16 / 89.02%±2.16	93.35%±2.10 / 94.89%±1.28	96.26%±1.83 / **97.07%±1.78**
CMU	91.6%±2.9	95.6%±1.7	94.8%±2.6	99.1%±0.6	98.0%±0.7	98.61%±1.3	90.24%±3.13 / 91.70%±2.86	94.31%±2.27 / 95.14%±2.92	97.89%±2.03 / **98.88%±0.42**
UWA	91.5%±3.07	-	-	98.2%±1.2	-	98.28%±1.05	89.57%±2.87 / 90.96%±1.74	91.20%±2.09 / 92.32%±2.54	97.11%±1.16 / **98.37%±1.13**

Table 5.1 shows the improvement provided by the proposed fusion method comparing to alternative smoothing techniques (averaging and energy). Proposed band fusion method using Max Pooling has higher recognition accuracy comparing to average (mean) of pixels and energy (mean squared value) of pixels along spectral axis since it keeps the maximum valued pixels which is called as most activated presence of a feature.

In this contribution DWT is employed for feature extraction along spectral dimension iteratively which after *n* level decomposition asymptotically has a computational complexity of $\mathcal{O}(NlogN)$ where $N$is an input's dimension. It is same/less comparing to mentioned state of the art methods in Table 5.1. In [13], they employed 3D-Gabor wavelet to extract features of hyperspectral face images which has the same computational complexity of $\mathcal{O}(N^3)$. In [15-16], features are extracted by 2D Fast Fourier Transform (FFT) from the log-polar images hence the computational complexity of FFT is $\mathcal{O}(NlogN)$. In [26], the first and second order statistics of each cubelets (section 2.4.2) are computed by the mean vector and covariance of 2D matrix which covariance matrix computation has a computational complexity of $\mathcal{O}(N^2)$.

5.6 Conclusion

In this chapter, we propose DWT based method to fuse spectral information by applying over each pixel along the spectral axis on hyperspectral image cubes. The DWT is applied consecutively to each vector of pixels extracted from the low frequency component from the previous decomposition. This process is iteratively repeated until the spectral vector for each pixel is decimated to a single pixel transforming the 3D input spectral image cube into a 2D output image. Transformed 2D images go through principal component analysis (PCA) for dimensionality

reduction generating the feature vectors to be classified by using *k*-NN and CRC classifiers. Experimental results using the proposed method show that, the performance in terms of accuracy outperforms state-of-the art methods using PolyU-HSFD, CMU-HSFD and UWA-HSFD databases.

Chapter 6

CONCLUSION AND FUTURE WORK

In this thesis, we have introduced novel methods for feature extraction of facial hyperspectral image classification to improve the recognition rate while using spatio-spectral information simultaneously.

6.1 Conclusion

3D-DWT is applied to whole hyperspectral image cube to extract features from the subbands generated by discrete wavelet decomposition in three different approaches. Three approaches included 3D-subband energy (3D-SE), 3D-subband overlapping cube (3D-SOC) and 3D-global energy (3D-GE). Each approach extracted different feature vector which contains the energy values calculated from different wavelet subbands at different levels of decomposition. Generated feature vectors by three different approaches went through a classifier to complete the face recognition task.

Furthermore, we proposed a novel method to fuse spectral information into a single 2D image by applying band-specific signal to noise ratio (SNR) weight. These weights are calculated based on band-specific SNR values to be assigned to each specific band for generating a single 2D face image. Hence, each pixel along spectral axis is fused to a single pixel resulting a 2D output face image for each 3D hyperspectral face cube. For dimensionality reduction, 2D output images went through principal component analysis (PCA) and face recognition was performed with the help of a classifier.

Finally, we employed discrete wavelet transform (DWT) to fuse spectral information by applying it to each pixel along the spectral axis on each hyperspectral image cube. The DWT was applied consecutively to each spectral vector of pixels extracted from the low frequency component from the previous decomposition. This process was iteratively repeated until the spectral vector for each pixel is decimated to a single pixel transforming the 3D input spectral image cube into a 2D output image. PCA was applied to 2D output images for dimensionality reduction and went through a classifier to perform a face recognition.

Some limitations should be noted. First, the methods are not robust for pose estimation, variance in orientation and face expression since all three standard hyperspectral databases are captured without considering mentioned limitations. Second, facial occlusion, such as sunglasses, scarf, mask etc., is one critical factor that affects the performance of face recognition and the proposed methods in this study are not robust to it.

The first contribution has a high computational complexity of $\mathcal{O}(N^3)$, second contribution's computational complexity is $\mathcal{O}(NlogN)$ which is low and the last contribution has a moderate computational complexity of $\mathcal{O}(N^2)$ comparing to some of the state of the art.

The experimental results revealed that recognition accuracy of all proposed methods by using standard hyperspectral databases outperform alternative hyperspectral face recognition of the state-of-the-art methods.

6.2 Future Work

Recently, deep Convolutional Neural Network (CNN) architectures have been introduced for the classification of images including faces. Many specialized CNN frameworks including AlexNet [59], GoogLeNet [60] and ResNet [61] have been introduced as pre-trained CNNSs with millions of samples for object classification. There are three major prospects involving CNNs. The first future work involves transfer learning, where pre-trained network is utilized to adapt to a new network by adjusting the parameters of the network to perform classification on Hyperspectral images. Second work can involve separate CNN pipelines for each/selected hyperspectral band is to be fused [62] for final classification. Third work can be 3D CNN [63] due to the 3D nature of Hyperspectral images. 3D Hyperspectral image cubes can be used to train dedicated 3D

REFERENCES

[1] Barnouti, N., Al-Dabbagh, S. & Al-Mansour, W. (2016). Face Recognition: A Literature Review. *International Journal of Applied Information Systems*, 11(4), 21-31.

[2] Zhao, W., Chellappa, R., Phillips, P. J. & Rosenfeld, A. (2003). Face recognition: A literature survey. *ACM Computing Surveys*, 35 (4), 399-458.

[3] Vartak P. & Bharadi, V. (2015). Hyperspectral face recognition by texture feature extraction using hybrid wavelets type I & II and kekre wavelet transform. *International Conference on Computing Control and Automation*, 721-727.

[4] PolyU-HSFD http://www.comp.polyu.edu.hk/~biometrics/hyper_face.htm, accessed 13 November 2015.

[5] CMU-HSFD. http://www.consortium.ri.cmu.edu/hsagree/index.cgi, accessed 16 May 2016.

[6] Khan, Z., Shafait, F. & Mian, A. (2015). Joint group sparse PCA for compressed hyperspectral imaging. *IEEE on Image Processing*, 24(12), 4934-4942.

[7] Chein, C. (2003). Hyperspectral Imaging: Techniques for Spectral Detection and Classification. *Springer Science & Business Media*. ISBN 978-1-4419-9170-6, 15-35.

[8] Pan, Z., Healey, G. & Tromberg, B. (2009). Comparison of spectral-only and spectral/spatial face recognition for personal identity verification. *EURASIP Journal on Advances in Signal Processing*, 2009 (1), 1-6.

[9] Di, W., Zhang, L., Zhang, D. & Pan, Q. (2010). Studied on hyperspectral face recognition in visible spectrum with feature band selection. *IEEE Transaction on Systems, Man, and Cybernetics, Part A*, 40(6), 1354-1361.

[10] Liang, J., Zhou, J. & Gao, Y. (2015). 3D local derivative pattern for hyperspectral face recognition. 11th *IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*. doi: 10.1109/FG.2015.7163115 .

[11] Uzair, M., Mahmood, A. & Mian, A. (2013). Hyperspectral face recognition using 3D-DCT and partial least squares. *BMVC*, 1-10.

[12] Pan, Z., Healy, G., Prasad, M. & Tromberg, B. (2003). Face recognition in hyperspectral images. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 25(12), 1552-1560.

[13] Shen, L. & Zheng, S. (2012). Hyperspectral face recognition using 3d Gabor wavelets. *Int. Conf. on Pattern Recognition*, pp. 1127-1137.

[14] Robila, S. (2008). Towards Hyperspectral Face Recognition. *Image Processing: Algorithms and Systems VI, 68120X*, doi: 10.1117/12.765268.

[15] Chen, G., Sun, W. & Xie, W. (2017). Hyperspectral face recognition with log-polar Fourier features and collaborative representation based voting classifiers. *IET Biometrics*, 6 (1), 36-42.

[16] Chen, G., Li, Ch. & Sun, W. (2017). Hyperspectral face recognition via feature extraction and CRC-based classifier. *IET Image Processing*, 11(4), 266-272.

[17] Ahonen, T., Hadid, A. & Pietikainen, M. (2004). Face recognition with local binary pattern. *Proc. of the 8th European Conf. on Computer Vision*, pp. 469-481.

[18] Dalal, N. & Triggs, B. (2005). Histograms of oriented gradients for human detection. *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, 1, 886–893. doi: 10.1109/CVPR.2005.177.

[19] Andersson, F. (2005). Fast inversion of the radon transform using log-polar coordinates and partial back-projections. *SIAM J. Appl.*, 65, 818–837.

[20] Štruc, V. & Pavešic, N. (2010). The complete Gabor-fisher classifier for robust face recognition. *EURASIP Adv. Signal Process.*, 2010**,** 1-26.

[21] Štruc, V. & Pavešic, N. (2009). Gabor-based kernel partial-least-squares discrimination features for face recognition. *Informatica (Vilnius)*, 20 (1), 115–138.

[22] Khotanzad, A. & Hong, Y.H. (1990). Invariant image recognition by Zernike moments. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12 (5), 489– 497.

[23] Zernike, F. (1934). Physica. *Elsevier,* vol. 1, pp. 689–704.

[24] Ahmed, N., Natarajan, T. & Rao, K.R. (1974). Discrete Cosine Transform. *IEEE Transactions on Computers*, C-23(1), 90-93.

[25] Khayam, S. A. (2003).The Discrete Cosine Transform (DCT): Theory and application. *Technical Report Michigan State University*.

[26] Uzair, M., Mahmood, A. & Mian, A. (2015). Hyperspectral face recognition with spatiospectral information fusion and PLS regression. *IEEE Transactions on Image Processing*, 24(3), 1127-1137.

[27] Zala, A. & Changela, A. (2014). 3D discrete wavelet transform VLSI architecture for image processing. *International Journal of Applied Engineering Research (IJAER)*, 2(3), 1826-1831.

[28] Tripathy, M., Sachdeva, K. & Talhi, R. (2009). 3D Discrete Wavelet Transform VLSI Architecture for Image Processing. *Proceedings, Progress in Electromagnetics research symposium, Moscow, Russia,* pp. 1569-1573.

[29] Weeks, M. & Bayoumi, M.A. (2002). Three-dimensional discrete wavelet transform architectures. *IEEE Transaction on Signal Processing*, 50(8), 2050-2063.

[30] Yoo, H.Y., Lee, K. & Kwon, B. (2007). Application of the 3D discrete wavelet transformation scheme to remotely sensed image classification. *Korean Journal of Remote Sensing*, 23(5), 355-363.

[31] Weeks, M.C. (1998). Architectures for the 3-D discrete wavelet transform. PhD thesis, The University of Southwestern Louisiana.

[32] Hsiang, S.T. & Woods, J.W. (2001). Embedded video coding using invertible motion compensated 3-D subband/wavelet filter bank. *Signal Process. Image Commun*., 16(8), 705–724.

[33] Challa, K., Krishna, P. & Rao, C. (2014). Design of a novel Architecture of 3-D Discrete Wavelet Transform for Image Processing through Video Compression. *International Conference on Communication and Signal Processing* (pp. 1092-1096). IEEE.

[34] Lustig, M., Donoho, D. & Pauly, J.M. (2007). Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn. Resonance Med*., 58(6), 1182-1195.

[35] Guo, X., Huang, X. & Zhang, L. (2014). Three-dimensional wavelet texture feature extraction and classification for multi/hyperspectral imagery. *IEEE Geoscience and Remote Sensing Letters*, 11(12), 2183-2187.

[36] Ghasemzadeh, A. & Demirel, H. (2016). Hyperspectral face recognition using 3D discrete wavelet transform. *Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA)*. doi: 10.1109/IPTA.2016.7821008.

[37] Ghasemzadeh, A. & Demirel, H. (2018). 3D discrete wavelet transform based feature extraction for hyperspectral face recognition. *IET Biometrics*, 7(1), 49-55.

[38] Ebrahimpour, H. & Kouzani, A. (2007). Face recognition using bagging KNN, *Int. Conf. Signal Processing and Communication Systems (ICSPCS)*, (pp. 17–19).

[39] Jose, J. P., Poornima, P. & Kumar, K. M. (2012). A novel method for color face recognition using kNN classifier. *International Conference on Computing, Communication and Applications.* IEEE. doi: 10.1109/ICCCA.2012.6179151.

[40] Kaur, M. & Dhriti. (2012). K-nearest neighbor classification approach for face *and fingerprint at feature level fusion. International Journal of Computer Applications*, 60(14). doi: 10.5120/9759-1517.

[41] Zhang, L., Yang, M. & Feng, X. (2011). Sparse representation or collaborative representation: which helps face recognition? *International Conference on Computer Vision,* (pp. 471-478). IEEE.

[42] Wandishin, M.S. & Mullen, S.J. (2009). Multiclass ROC analysis. *Weather Forecast,* 24(2), 530–547.4

[43] Polikar, R. (2001, January 12). The wavelet tutorial. Retrieved from http:// users.rowan.edu/~polikar/WAVELETS/WTpart1.html.

[44] Masayuki Tanka's Matlab code for face parts detection. Retrieved from http://www.mathworks.com/matlabcentral/fileexchange/36855-face-parts detection.

[45] Pentland, A. & Turk, M. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, vol. 3, 71-86.

[46] Shlens, J. (2009). A tutorial on principal component analysis. In *Center for Neural Science and Systems Neurobiology Laboratory, Salk Institute for Biological Studies La Jolla*, version 3.01, vol. 3.

[47] Lindsay, I. S. (2002). *A tutorial on principal component analysis.* Retrieved from http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components

[48] Sriovich, L. & Kirbly, M. (1990). Application of the Karhunen-Loeve procedure for the characterization of human faces. *IEEE Transaction PAMI,* vol. 12, 103-108.

[49] Chippy, J., Jacob, N. V., Renu, R. K., Sowmya, V. & Soman, K. P. (2017). Least square denoising in spectral domain for hyperspectral images. *7th International Conference on Advances in Computing & Communications, ICACC,* (pp. 399-406). Elsevier.

[50] Chang, D., Zhao, Y., Zheng, C. & Zhang, X. (2012). A genetic clustering algorithm using a message-based similarity measure. *Expert Systems with Applications*, vol. 39, 2194-2202.

[51] Ju, C. & Xu, C. (2013). A new collaborative recommendation approach based on users clustering using artificial bee colony algorithm. *The Scientific World Journal,* vol. 2013.

[52] Wu, Q., Qi, X., Fuller, E. & Zhang, C.Q. (2013). Follow the leader: A centrality guided clustering and its application to social network analysis. *The Scientific World Journal,* vol. 2013.

[53] Farhang, Y. (2017). Face extraction from image based on k-means clustering algorithms. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 8(9), 96-107.

[54] Rasti, B., Scheunders, P., Ghamisi, P., & Licciardi, G.: (2018). Noise Reduction in Hyperspectral Imagery: Overview and Application. *Remote Sens.,* 10, (3), pp. 482–510.

[55] Xu, L., Li, L., Wong, A. & Clausi, D. A. (2015). Hyperspectral Image Denoising Using a Spatial–Spectral Monte Carlo Sampling Approach. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8, (6), pp. 3025–3038.

[56] Daubechies wavelets are useful in noise removal. (2001, July). Retrieved from http://bearcave.com/misl/misl_tech/wavelets/haar.html.

[57] James S. Walker. (2008). *A Primer on Wavelets and Scientific Applications*. Boca Raton, Florida: CRC Press.

[58] Adão T., Hruška J. *et al.* (2017). Hyperspectral Imaging: A Review on UAV-Based Sensors, Data Processing and Applications for Agriculture and Forestry. *Remote Sensing Journal,* 9 (11), 1110.

[59] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D. Vanhoucke, V. and Rabinovich, A. (2015). Going deeper with convolutions. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 1-9.

[60] Szegedy, C., Vanhoucke, V., Ioffe, S., *et al.* (2016). Rethinking the inception architecture for computer vision. *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 2818–2826.

[61] Krizhevsky, A., Sutskever, I., Hinton, G.E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems,* pp. 1097–1105.

[62] Serte, S. and Demirel, H. (2019). Gabor wavelet-based deep learning for skin lesion classification. *Computers in Biology and Medicine,* vol.113.

[63] Sidra, R., Zahid, A., Unsang, P., *et al.* (2019). Age-invariant face recognition using gender specific 3D aging modeling. *Multimedia Tools and Applications,* 78 (17), pp 25163–25183.

[64] Feng, J, Guo, Q., Guan, Y., *et al.* (2018). 3D Face Recognition Method Based on Deep Convolutional Neural Network. *Smart Innovations in Communication and Computational Sciences,* pp. 123-130.

[65] Tobías, L., Ducournau, A., Rousseau, F., *et al.* (2016). Convolutional Neural Networks for object recognition on mobile devices: A case study. *23rd International Conference on Pattern Recognition (ICPR),* pp. 3530–3535.