ENSEMBLE OF MULTIPLE MATCHERS FOR FINGER VEIN RECOGNITION

SOH SIANG LOONG

UNIVERSITI SAINS MALAYSIA

2017

ENSEMBLE OF MULTIPLE MATCHERS FOR FINGER VEIN RECOGNITION

by

SOH SIANG LOONG

Thesis submitted in partial fulfilment of the requirements of the degree of Bachelor of Engineering (Electronic Engineering)

JUNE 2017

ACKNOWLEDGEMENTS

I would like to use this opportunity to express my gratitude to everyone who supported and guided me throughout this final year project.

Above all, I would like to thank my supervisor, Dr. Bakhtiar Affendi bin Rosdi for his guidance. He advised and guided me patiently until the project is completed. When I encountered problems, he is able to assists and gives opinion to provide me solutions for it. He also uses Slack as platform to discuss the project and the responses are very fast. He is also willing to sacrifice his time for proofreading my thesis and gives comments to improve it. Next, I would like to thank my examiner, Dr. Haidi bin Ibrahim for giving me feedback on my thesis.

Besides, I wish to thank my friends for giving opinions and helps me whenever I encountered problems. Without them, I would have hard time struggling in complete this project. Lastly, I want to thank my family members who supported me mentally and financially during the four years of studies for degree of electronic engineer in Universiti Sains Malaysia.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2D	Two Dimensional
BLPOC	Band-Limited Phase Only Correlation
CCD	Centroid Contour Distance
DFT	Discrete Fourier Transform
EER	Equal Error Rate
IDFT	Inverse Discrete Fourier Transform
LDA	Linear Discriminant Analysis
MGF	Modified Gaussian Filter
NIR	Near-Infrared
NN	Nearest Neighbour
PCA	Principal Component Analysis
POC	Phase-Only Correlation
RBF	Radial Basis Function
ROI	Region of Interest
SRC	Sparse Representation Classifier
SVM	Support Vector Machine
W	Width
WCCD	Width-Centroid Contour Distance

LIST OF SYMBOLS

a	mean
$A \times B$	patch size for spatial pyramid pooling
$A_F(u, v)$	amplitude of $F(u, v)$
$A_G(u, v)$	amplitude of $G(u, v)$
b	parameter in MGF
c	parameter in MGF
D	Euclidean distance
D_{lr}	direction for tracing point to move (left or right)
D_{ud}	direction for tracing point to move (up or down)
e	scores obtained from feature extractions
$e_i{}'$	normalized e
$e_i{}''$	modified $e_i{}'$
f	pooled feature from spatial pyramid pooling
$f(x, y)$	finger vein image 1
f_s	final score for fusion of algorithm
$F(u, v)$	DFT of $f(x, y)$
$g(x, y)$	finger vein image 2
$G(u, v)$	DFT of $g(x, y)$
$\overline{G(u, v)}$	conjugate of $G(u, v)$
$H(x, y)$	MGF formula
l	number of patches
K	RBF kernel's matrix

m	frequency range for BLPOC
$M \times N$	dimension of finger vein image for BLPOC
n	frequency range for BLPOC
N_3	a set of three neighbouring pixels which are the possible movement for tracking to move
N_8	a set of eight neighbouring pixels which are the possible movement for tracking to move
N_c	a set of pixels which are the possible movement for tracking to move
$N_r(x_c, y_c)$	neighbouring pixels of (x_c, y_c)
p	number of patches in pooling cell
p_{lr}	probability of selecting the three neighbouring pixels (left or right)
p_{ud}	probability of selecting the three neighbouring pixels (up or down)
q	band-limit factor
r	distance between (x_c, y_c) and cross section
$r_{fg}(u, v)$	original POC function of two images, IDFT of $R_{FG}(u, v)$
R_f	random set of coordinate
$R_{FG}(u, v)$	cross-phase spectrum of images
R_{nd}	uniform random number between 0 and n
s_g	normalized geometry score obtained from WCCD
s_v	normalized vein score obtained from BLPOC
t_x, t_y	translation parameters between images
T_c	locus-position table
$\overline{T_c}$	conjugate of T_c
v	standard deviation

V_l	constant to check if position of (x_c, y_c) should be updated
w_g	weighting factor for geometry
w_v	weighting factor for vein
W	width
x	row vector formed from patches
x_c, y_c	current tracking point
x_s, y_s	starting point for tracking
$X \times Y$	dimension of finger vein image
θ	angle between dark line and horizontal axis of (x_c, y_c)
$\theta_F(u, v)$	phase of $F(u, v)$
$\theta_G(u, v)$	phase of $G(u, v)$
γ	free parameter of RBF

PENGGABUNGAN KAEDAH-KAEDAH PEMADANAN BAGI PENGIKTIRAFAN SALURAN DARAH JARI

ABSTRAK

Sistem pengiktirafan biometrik merupakan sistem yang amat penting dalam pengenalan dan pengesahan individu. Penyelidikan bagi pengesahan saluran darah jari semakin popular disebabkan manfaat-manfaat yang diperoleh seperti kebersihan dan tidak boleh ditiru. Selain itu, pengesahan saluran darah jari juga dapat mengatasi keperluan masyarakat dan masalah kesihatan. Pelbagai kaedah pengekstrakan ciri telah dicadangkan oleh penyelidik, seperti pengesanan garis berulang kali, “Principal Component Analysis” (PCA), “Linear Discriminant Analysis” (LDA) dan Saluran Terhad Fasa Sahaja Korelasi (BLPOC). Kaedah-kaedah tersebut dikategorikan sebagai pengekstrakan ciri secara buatan tangan. Pengekstrakan ciri secara belajar masih tidak pernah digunakan dalam saluran darah jari. Jadi, pengumpulan data secara piramid digunakan dalam pengekstrakan ciri saluran darah jari. BLPOC digunakan sebagai pengekstrakan ciri secara buatan tangan. Skor-skor yang diperoleh akan digabungkan dengan skor-skor yang diperoleh dari pengumpulan data secara piramid dengan menggunakan Mesin Vektor Sokongan (SVM). Pangkalan data FV-USM yang diguna terdiri daripada 123 individu, imej saluran darah 4 jari bagi setiap individu ditangkap. Kadar Kesilapan Sama (EER) bagi pengumpulan data secara piramid adalah paling tinggi, sebanyak 4.368%, diikuti dengan BLPOC, 2.36% dan SVM mempunyai EER yang paling rendah 0.1348%. Konklusinya, gabungan antara pengekstrakan ciri secara belajar dan secara buatan tangan menunjukkan keputusan yang lebih baik dibandingkan dengan pemadanan ciri secara tunggal.

ENSEMBLE OF MULTIPLE MATCHERS FOR FINGER VEIN RECOGNITION

ABSTRACT

Biometrics recognition system is important in identification and verification of an individual. Recently, the research on finger vein verification becomes popular due to the benefits such as hygiene and cannot be duplicated. Finger vein verification is also able to overcome community needs and health problems. Various feature extraction methods were proposed by researchers, such as repeated line tracking, Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Band-Limited Phase-Only Correlation (BLPOC. These methods are considered as hand-crafted feature extraction method. Learned feature extraction has not been used in finger vein verification yet. Hence, spatial pyramid pooling method is developed as learned feature extraction for finger vein verification. BLPOC is used as hand-crafted feature extraction which the scores obtained will be then fused together with the scores obtained from spatial pyramid pooling by using Support Vector Machine (SVM). The database used is FV-USM based on 123 individuals with 4 fingers each. In the result obtained, spatial pyramid pooling shows the highest EER, 4.368%, followed by BLPOC, 2.36% and the lowest is SVM, 0.1348%. As conclusion, fusion of learned feature and hand-crafted feature shows the best performance as compared to single feature matching.

CHAPTER 1

INTRODUCTION

1.1 Background

Biometrics are useful in many ways, especially for verification and identification. In verification, an individual has to declare himself as a specific person and the system will check if his or her biometric matches with the record in database. Verification process does not access the whole database as they would only have to go through 1-to-1 matching test for the specific person's record. As for identification, it needs to read unknown biometric and access the database for 1-to-n matching test to track his or her information (Mayhew, 2012).

The common physical characteristics used in biometrics recognition are face, fingerprints and hand geometry. However, some of the characteristics are easily exposed to others and can be duplicated (Rosdi, Shing, & Suandi, 2011). Moreover, there are some factors that might hinder the effectiveness of fingerprint verification such as sweaty palms, as their fingerprint patterns may not be clear (Mohd Asaari, Suandi, & Rosdi, 2014). Drahansky, Dolezel, Urbanek, Brezinova, & Kim (2012) mentioned that there are several types of skin disease attack hand palms and fingertips, which brings distortion to their fingerprint and leads to problem in fingerprint recognition technology. Thus, finger vein verification is now being introduced (Kono, Euki, & Umemura, 2002). Finger vein verification has few benefits, such as live-body identification, non-contact (hygiene), hard to duplicate, and low failure rate (Gupta & Gupta, 2015). By using finger vein verification, the problems faced by community as mentioned above will be solved and does not affect human's health.

Figure 1.1 shows the block diagram of the procedure for finger vein verification. The procedure is divided into five main stages: image acquisition, pre-processing, feature extraction, feature matching and analyse result (Mohd Asaari et al., 2014; Wu & Liu, 2011). In image acquisition, the images of vein pattern are captured by infrared CCD camera using Near-Infrared (NIR) light as the light source. The range of wavelength for NIR light is 700 nm to 1000 nm, however 850 nm is commonly used in capturing finger vein images. Then, the raw image will undergo pre-processing to extract the Region of Interest (ROI). After the finger vein images are processed, they will undergo feature extraction, which will be used in matching test.

Figure 1.1: Procedure for finger vein verification

There are two types of feature extraction, namely hand-crafted and learned features extraction. Hand-crafted features use manually predefined algorithms based on the expert knowledge to extract the important information such as finger vein. On the other hand, learned features are derived from an image dataset by training procedure in order to perform verification process (Antipov, Berrani, Ruchaud, & Dugelay, 2015).

The common methods used in feature extraction for finger vein are Principal Component Analysis (PCA) (Shen, Shen, Zhou, Yang, & Shen, 2014; Wu & Liu, 2011; Yang & Zhang, 2012) and repeated line tracking (Miura, Nagasaka, & Miyatake, 2004). Besides, Linear Discriminant Analysis (LDA) (Wu & Liu, 2011) and Band-Limited Phase-Only Correlation (Mohd Asaari et al., 2014) are also the methods for feature extraction.

However, there are some researches mentioned that fusion of several algorithms showed better result than single algorithm. There are several methods for score-level fusion. In Karanwal, Kumar, & Maurya (2010), the results of max rule, min rule and product rule in score-level fusion is discussed. Besides, the research by Mohd Asaari et al. (2014) which fused the scores obtained from two different algorithms by using weighted sum rule shows that the accuracy from fusion is higher than unimodal. Therefore, this research is carried out to extract feature obtain from learning-based and hand-crafted methods, then fused together using score-level fusion.

1.2 Problem Statements

Feature extraction is very important stage for finger vein verification. Researches have been carried out to improve the verification accuracy. The feature extraction methods that are proposed for finger vein are considered as hand-crafted features. However, there are limitations in hand-crafted features. For example, Principal Component Analysis (PCA) extracts the features by using global based approach. However, feature extracted by PCA ignores local information (Xi, Yang, Yin, & Meng, 2013). Besides, the method proposed by Miura, Nagasaka, & Miyatake (2004) required predefined probability and might miss some part of image as the initial tracking point is random.

Besides, there are also limitations for single feature matchers. Band-Limited Phase Only Correlation (BLPOC) (Rosdi et al., 2011) only considers phase information and ignores the magnitude information. Thus, score-level fusion is carried out in Mohd Asaari et al. (2014) by fusing vein score and geometry score using weighted sum rule to overcome the problem.

1.3 Objectives

The objectives to be achieved in this project are as follow:

1. To improve accuracy of finger vein verification by fusing both hand-crafted feature and learned feature.
2. To evaluate the performance of the proposed algorithm.

1.4 Scopes of Project

This project aims to improve the recognition accuracy by ensemble multiple matchers. The scope of this project focuses on feature extraction and matching. Image acquisition and ROI extraction is not covered. The finger vein images used are extracted ROI images. These images are in 8-bit-per-pixel grayscale digital images, with dimension of 100×300 pixels.

The total number of images used in this project is 5904 ($2 \times 123 \times 4 \times 6$, images taken in 2 sessions, 123 individuals, 4 fingers each, 6 images taken from each finger). Verification is used for this project to check if he or she is the person specified. Besides, this project used local image processing method which process the image by patches instead of global image processing method which process the whole image.

1.5 Thesis Organization

This report consists of five chapters. Chapter 1 provides a brief introduction to the project, this includes the overview of finger vein recognition, problem statement, objectives, scope of project, and organization of report.

Chapter 2 reviews about the past researches on the finger vein feature extraction and image classifier. This chapter also discusses the pros and cons of the methods reviewed. Chapter 3 shows the details of spatial pyramid pooling for feature extraction. After the features are extracted, the data will be used for score calculation. The result obtained will be compared with previous existing methods.

Chapter 4 discusses on the result of the research. The parameters used for spatial pyramid pooling and Support Vector Machine (SVM) are determined. Besides, the result obtained from spatial pyramid pooling, BLPOC and fusion of the methods using SVM are analysed. Lastly, Chapter 5 concludes the outcome of this research and discusses about the future work which can improve this research.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Biometrics are widely used in verification and identification. Recently, finger vein has become more popular in research. There are researches on various methods for finger vein feature extraction.

In order to improve the accuracy of finger vein verification, fusion of algorithms is introduced. By using this method, the efficiency and accuracy can be improved as there are pros and cons for every method proposed by researchers.

In this chapter, the methods proposed for feature extractions by previous researches are discussed. Besides, some of the fusion methods are discussed too. Lastly, the performance evaluation method will be discussed in this chapter too.

2.2 Feature Extraction for Finger Vein Verification

There are two types of feature extraction, which are hand-crafted and learning based features. Most of the methods for finger vein's feature extraction methods proposed by researchers are categorized as hand-crafted features extraction.

This section discussed on methods for hand-crafted feature extraction. The methods discussed are repeated line tracking (Miura et al., 2004), Principal Component Analysis (PCA) (Shen et al., 2014; Wu & Liu, 2011; Yang & Zhang, 2012), Linear Discriminant Analysis (LDA) (Wu & Liu, 2011), Phase-Only Correlation (POC), and

Band-Limited Phase-Only Correlation (BLPOC) (Ito, Member, Nakajima, & Kobayashi, 2004; Mohd Asaari et al., 2014; Rosdi et al., 2011).

2.2.1 Repeated Line Tracking

Miura et al. (2004) proposed repeated tracking of dark lines in images to keep the extraction robust against irregular shading and noise. The overall steps involved is shown in Figure 2.1.

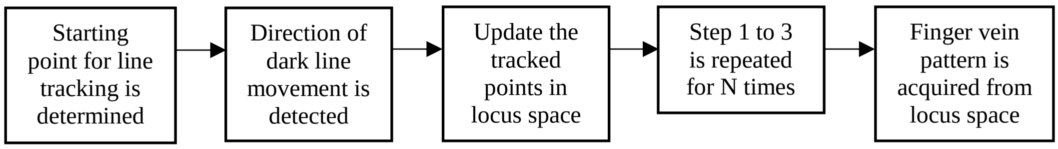


Figure 2.1: Procedure for repeated line tracking method

First, the starting point for line tracking is determined. The starting point is selected randomly from random set of coordinate R_f and labelled as (x_s, y_s). The value of current tracking point is (x_c, y_c), which means that the initial value of (x_c, y_c) is equivalent to (x_s, y_s). Then, the direction for the tracing point to move left and right D_{lr}, or up and down D_{ud} is determined, with the function as shown in (2.1) and (2.2) (Miura et al., 2004):

$$D_{lr} = \begin{cases} (1, 0), & \text{if } R_{nd}(2) < 1 \\ (-1, 0), & \text{otherwise} \end{cases} \tag{2.1}$$

$$D_{ud} = \begin{cases} (0, 1), & \text{if } R_{nd}(2) < 1 \\ (0, -1), & \text{otherwise} \end{cases} \tag{2.2}$$

where $R_{nd}(n)$ is random integer between 0 and n.

After the initial tracking point and direction is determined, the dark-line direction is detected for the tracking point to move. A locus-position table T_c is created to store the position of the tracking point moved. Next, a set of pixels N_c which are the possible movements of the tracking point are determined. The previous tracking point must not be included. N_c is determined as (2.3) (Miura et al., 2004):

$$N_c = \overline{T_c} \cap R_f \cap N_r(x_c, y_c) \tag{2.3}$$

where $\overline{T_c}$ is the conjugate of T_c and $N_r(x_c, y_c)$ is the set of neighbouring pixels of (x_c, y_c), as shown in (2.4) (Miura et al., 2004):

$$N_r = \begin{cases} N_3(D_{lr)}(x_c, y_c), & \text{if } R_{nd}(100) < p_{lr} \\ N_3(D_{ud)}(x_c, y_c), & \text{if } p_{lr} + 1 \le R_{nd}(100) < p_{lr} + p_{ud} \\ N_8(x_c, y_c), & \text{if } p_{lr} + p_{ud} + 1 \le R_{nd}(100) \end{cases} \tag{2.4}$$

where p_{lr} and p_{ud} are the probability of selecting the three neighbouring pixels, $N_8(x_c, y_c)$ is the set of eight neighbouring pixels of (x_c, y_c) and $N_3(D)(x_c, y_c)$ is the set of three neighbouring pixels of (x_c, y_c),.which is determined using (2.5) (Miura et al., 2004):

$$N_3(D)(x, y) = \{(D_x + x, D_y + y), (D_x - D_y + x, D_y - D_x + y), \tag{2.5}$$

$$(D_x + D_y + x, D_y + D_x + y)\}$$

where D (defined as (D_x, D_y)) is the moving direction to determine N_3.

Next, the pixel which the current tracking point should move is detected using line-evaluation function (2.6) (Miura et al., 2004). The variables used for this function is illustrated in Figure 2.2. After the pixel movement is confirmed, the data is recorded in locus-position table T_c. Position for current tracking point (x_c, y_c) is added to the table. After that, (x_c, y_c) is updated to (x_i, y_i) if the line-evaluation function, V_l is positive and maximum. The steps are repeating until V_l is negative or zero, as (x_c, y_c) is not on dark line. After that, the number of times points in the locus has been tracked $T_r(x, y)$ is updated. The steps from assigning initial tracking point until T_c is updated is repeated for N times. Finally, the feature extraction is done. The result of finger-vein extraction is shown in Figure 2.3.

$$V_l = \max_{(x_i,y_i)\in N_c} \left\{ F\left(x_c + r\cos\theta_i - \frac{W}{2}\sin\theta_i, y_c + r\sin\theta_i + \frac{W}{2}\cos\theta_i\right) + F\left(x_c + r\cos\theta_i + \frac{W}{2}\sin\theta_i, y_c + r\sin\theta_i - \frac{W}{2}\cos\theta_i\right) - 2F(x_c + r\cos\theta_i\,, y_c + r\sin\theta_i) \right\} \tag{2.6}$$

where r is the distance between (x_c, y_c) and the cross section, θ_i is the angle between $(x_c, y_c) - (x_c + 1, y_c)$ and $(x_c, y_c) - (x_i, y_i)$, and W is the width of cross section.

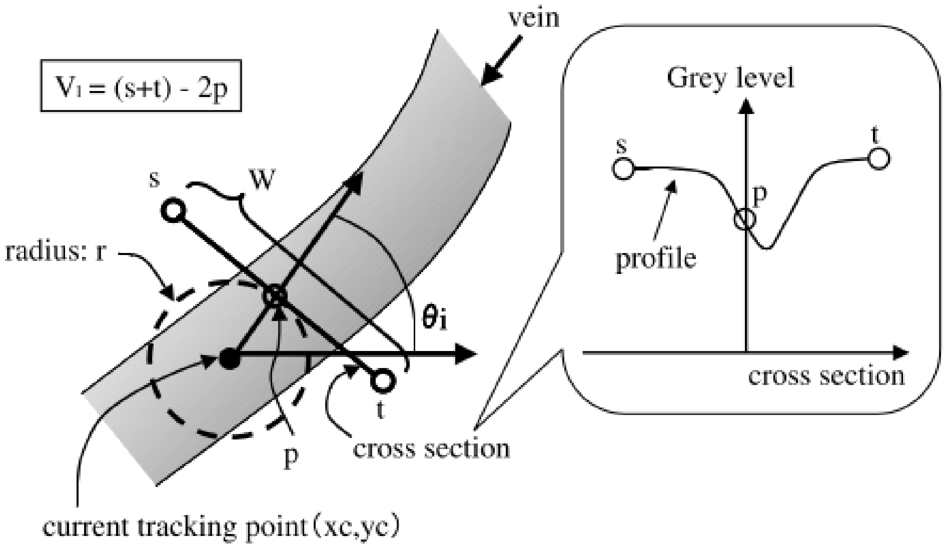


Figure 2.2: Dark-line detection (Miura et al., 2004)

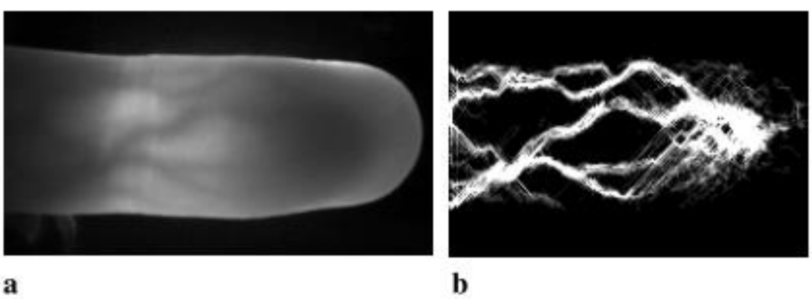


Figure 2.3: Effectiveness of finger-vein-pattern extraction (a) Infrared image; (b) Value distribution in the locus space (Miura et al., 2004)

However, repeated line tracking method required predefined probability in the horizontal and vertical orientation to choose the direction for the tracking point to move. Besides, the initial tracking point is selected randomly, which might miss some part of the finger vein image. This method shows the result with 0.145% error for a mismatch ratio of 37.6%.

2.2.2 Principal Component Analysis (PCA)

According to Jolliffe (2002), the main idea for PCA is to simplify multidimensional to lower dimensional. Eigenvectors and eigenvalues are obtained from the set of data. The direction of line is called as eigenvector while the variance of the data in that direction is called as eigenvalue. The dimension is then reduced by selecting the eigenvector with highest eigenvalue.

Wu & Liu (2011) proposed feature extraction on finger vein by applying PCA to extract global features. PCA is applied to reduce the dimensionality by discarding the eigenvectors with low eigenvalues as major information of the data is carried in eigenvectors with higher eigenvalues. Sample images for original image and after feature extraction is shown in Figure 2.4.

PCA is also used by other researchers. Yang & Zhang (2012) applied PCA after the image is enhanced to deal with small sample problem by reducing its dimensionality. Other biometrics such as face recognition also apply PCA. Shen et al. (2014) applied PCA on face image to reduce noise and also make sure the final pooled feature’s dimension is not too large.

However, Xi et al. (2013) mentioned that the method proposed may ignore some local detailed information as the PCA extracted global features and directly apply on image classifier. The error when apply only PCA is high, which is around 14% in Wu & Liu (2011). Thus, the feature extracted from PCA must be further processed by other methods to reduce error.

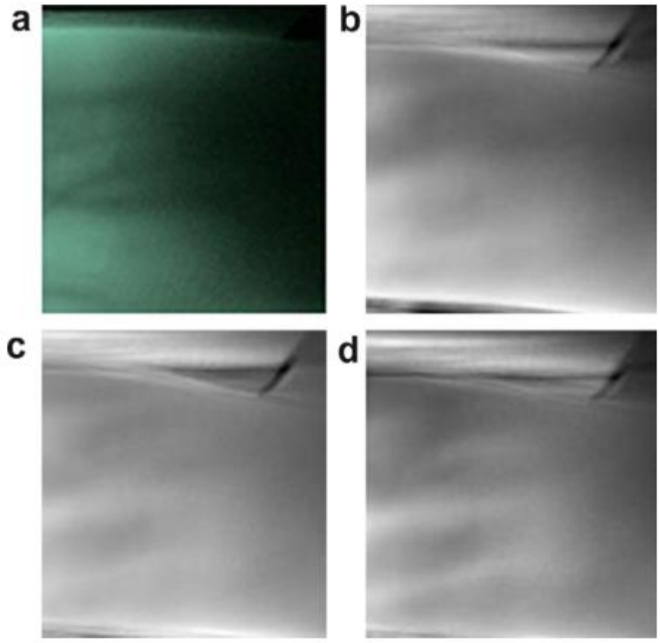


Figure 2.4: Sample feature images (a) original image; (b) four-feature image; (c) five feature image; (d) six-feature image (Wu & Liu, 2011)

2.2.3 Linear Discriminant Analysis (LDA)

LDA is a linear transformation technique used for reducing the dimension of an image, which has the common function as PCA. However, there are slight differences between PCA and LDA. PCA obtains the eigenvectors as whole, ignoring the class labels and variance in dataset; LDA obtains the eigenvectors which have the maximum separation between multiple classes. It is common for PCA and LDA to be used together (Raschka, 2014).

Wu & Liu (2011) carried out research on both PCA and PCA+LDA feature extraction process. The result is shown in Figure 2.5. PCA+LDA method is able to extract the discriminative features more precisely. However, LDA is very similar to PCA, thus it may also ignore some local detailed information.

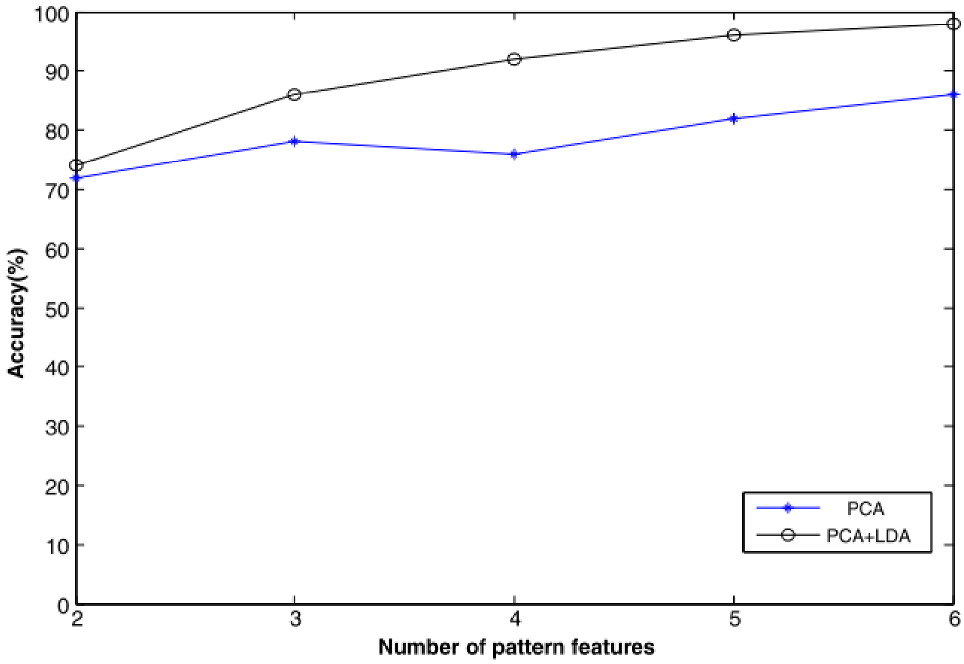


Figure 2.5: Accuracy rate as a function of pattern features in PCA and PCA+LDA (Wu & Liu, 2011)

According to Guan, Wang, & Yang (2011), there is a problem in LDA, which is LDA needs to transform image matrix into high dimension vector as PCA. Therefore, Two Dimensional LDA is introduced to avoid the similar problem. However, the result from this research shows that the error is quite high, with 4.55%.

2.2.4 Phase-Only Correlation (POC) and Band-Limited POC (BLPOC)

POC is used to evaluate the translation parameters between two finger vein images from the same individual. The translation parameters (t_x, t_y) between the images can be estimated from the peak of the graph obtain from POC function. The common regions are extracted after the images are aligned based on (t_x, t_y). If the value of t_x and t_y exceed 20 pixels and 10 pixels respectively, the images will not be translated as they are most probably from two different individuals (Rosdi et al., 2011).

Ito et al. (2004) and Mohd Asaari et al. (2014) explained the definition for POC function. Two $X \times Y$ images $f(x, y)$ and $g(x, y)$ are considered. 2D Discrete Fourier Transform (DFT) is applied on both $f(x, y)$ and $g(x, y)$, denoted as $F(u, v)$ and $G(u, v)$ respectively. $F(u, v)$ and $G(u, v)$is given as (2.7) and (2.8) respectively (Mohd Asaari et al., 2014):

$$F(u,v) = \sum_{x=0}^{X}\sum_{y=0}^{Y} f(x,y) e^{\frac{-j2\pi ux}{X}} e^{\frac{-j2\pi uy}{Y}} \tag{2.7}$$

$$= A_F(u,v) e^{j\theta_F(u,v)}$$

$$G(u,v) = \sum_{x=0}^{X}\sum_{y=0}^{Y} g(x,y) e^{\frac{-j2\pi ux}{X}} e^{\frac{-j2\pi uy}{Y}} \tag{2.8}$$

$$= A_G(u,v) e^{j\theta_G(u,v)}$$

where $A_F(u, v)$ and $A_G(u, v)$ are the amplitude and $\theta_F(u, v)$ and $\theta_G(u, v)$ are the phases. The cross-phase spectrum of the images is obtained using (2.9) (Mohd Asaari et al., 2014):

$$R_{FG}(u,v) = \frac{F(u,v)\overline{G(u,v)}}{\left|F(u,v)\overline{G(u,v)}\right|} \tag{2.9}$$

$$= e^{j\theta(u,v)}$$

where $\overline{G(u,v)}$ is the complex conjugate of $G(u, v)$ and $\theta(u, v)$ is the phase difference, obtained from $\theta_F(u, v) - \theta_G(u, v)$. 2D Inverse Discrete Fourier Transform (IDFT) is applied on $R_{FG}(u, v)$ to obtain the original POC function of these two images $r_{fg}(x, y)$. The equation is shown in (2.10) (Mohd Asaari et al., 2014):

$$r_{fg}(x,y) = \frac{1}{XY}\sum_{u=0}^{X}\sum_{v=0}^{Y} R_{FG}(u,v)\, e^{\frac{-j2\pi ux}{X}} e^{\frac{-j2\pi uy}{Y}} \tag{2.10}$$

However, POC function gives distinct sharp peak for identical vein images and low peak for different images. Low peak indicates the frequency components have low power and frequency higher than the dominant frequency spectrum. These frequency components usually contain less meaningful information. Band-Limited Phase-Only Correlation (BLPOC) is introduced to filter and eliminate these meaningless high-frequency components (Ito et al., 2004; Mohd Asaari et al., 2014).

In BLPOC, new size for frequency spectrum is given as M and N. The size is determined by a band-limit factor q. Thus the BLPOC function is given as (2.11) (Mohd Asaari et al., 2014):

$$r_{fg}^{mn}(x,y) = \frac{1}{MN}\sum_{u=0}^{M}\sum_{v=0}^{N} R_{FG}(u,v)\, e^{\frac{-j2\pi um}{M}} e^{\frac{-j2\pi un}{N}} \tag{2.11}$$

The error shown in Mohd Asaari et al. (2014) is 2.34%. However, BLPOC only takes the phase into account, neglecting the magnitude of the vein images.

2.3 Learned Features

In feature extraction for finger vein verification, hand-crafted features extraction is commonly used, such, PCA and LDA (Wu & Liu, 2011), BLPOC (Mohd Asaari et al., 2014), multi-scale matched filtering with line tracking (Gupta & Gupta, 2015) and efficient local binary pattern (Liu & Kim, 2016). Predefined algorithms based on the expert knowledge to extract the important information is required in hand-crafted features extraction while learned features are derived directly from image dataset by training procedure. Tan, Bhanu, & Lin (2005) mentioned that learned features are extracted from learning composite operators based on primitive features automatically. Useful unconventional features which are difficult for human to analyse and visualise can be extracted.

Other than biometrics, learned feature also applied in other human characteristics such as gender. Antipov et al. (2015) carried out research in comparing learned and hand-crafted features for pedestrian gender recognition. The result obtain through the research shows that learned feature has better performance compared to hand-crafted features.

Spatial pyramid pooling is used in extracting learned features. This is a method which divide an image, and extract the local features on them into areas from finer to coarser levels. By using spatial pyramid pooling, a fixed-length output will be generated regardless the size for input. Besides, the flexibility of input scales allows spatial pyramid pooling to be able to pool features extracted at different scales (He, Zhang, Ren, & Sun, 2015).

Shen et al. (2014) proposed a method to compute the facial features by performing pooling directly on the local patches based on spatial pyramid pooling. There are two popular pooling methods used in their algorithms, which are average pooling (2.12) and max pooling (2.13) (Shen et al., 2014). The average-operation or max-operation can be performed easily on normalized pixel of the local patches.

$$f_i = \sum_j x_i^{(j)}/p \quad \text{(average pooling)} \tag{2.12}$$

$$f_i = \max_j x_i^{(j)} \quad \text{(max pooling)} \tag{2.13}$$

where $x_i^{(j)}$ is the ith element of the jth local patch in the current pooling cell, p is the number of patches in pooling cell and $f = [f_1, \ldots, f_i, \ldots]$ is the pooled feature for that cell. The result from research shows average pooling is better than max pooling when the pooling levels are below six as max pooling discarded too much information. However, if the pooling levels are above six, max pooling shows higher accuracy than average pooling (Shen et al., 2014).

Shen et al. (2014) mentioned that this method is much simpler and obtains better result. This method can also be applied in many other vision tasks. However, learned feature extraction has not been used in finger vein images yet.

2.4 Fusion of Algorithms

Fusion of algorithms allows us to overcome some limitation by previous researches. There are few researches carried out on fusion of algorithms to improve the performance of finger vein verification. Yang & Zhang (2010) proposed feature-level fusion by fusing global and local features, Lu, Yoon, & Park (2013) proposed score-level fusion of Gabor features, Mohd Asaari et al. (2014) carried out score-level fusion for algorithms using weighted sum rule and Oloyede & Hancke (2016) fused multimodal biometric by several level of fusion strategy, such as sensor-level, feature-level, score-level and decision-level fusion.

Mohd Asaari et al. (2014) mentioned that the geometry feature in method proposed by Kang, Park, Yoo, & Kim (2011) can be affected by the rotation and translation of finger. Thus, Mohd Asaari et al. (2014) proposed fusion of finger thickness or width (W) (Kang & Park, 2009) and Centroid Contour Distance (CCD) (Fotopoulou, Laskaris, Economou, & Fotopoulos, 2013), forming a new geometrical feature called Width-Centroid Contour Distance (WCCD). Recognition accuracy increased significantly by this fusion. Besides, Mohd Asaari et al. (2014) also proposed to fuse the normalized vein score s_v obtained from Band-Limited Phase-Only Correlation (BLPOC) function and normalized geometry score s_g obtained from WCCD technique to generate final matching scores. Weighted sum rule is used to fuse the normalized scores (2.14) (Mohd Asaari et al., 2014).

$$f_s = w_v s_v + w_g s_g \tag{2.14}$$

where f_s is the final score, w_v and w_v are the weighting factors for vein and geometry respectively. By using fusion strategy, the limitation where BLPOC only takes the phase into account can be overcome. The result in this research shows a great improvement in performance. The Equal Error Rate (EER) for WCCD is 7.01% and BLPOC is 2.34%. After fusion, the EER is reduced to 1.78%.

Support Vector Machine (SVM) is also commonly used in score-level fusion. There are several researches carried out using score-level fusion on biometrics recognition. Park & Park (2007) fused the scores obtained from two different filters while Gawande et al. (2013) and Kumar & Devi (2014) fused the scores obtained from iris images and fingerprint images. Before fusion, the scores must be normalized, labelled with +1 for genuine scores and -1 for imposter scores. Then, the scores are separated into training score and testing score as input for SVM. The general flow for SVM is shown in Figure 2.6.

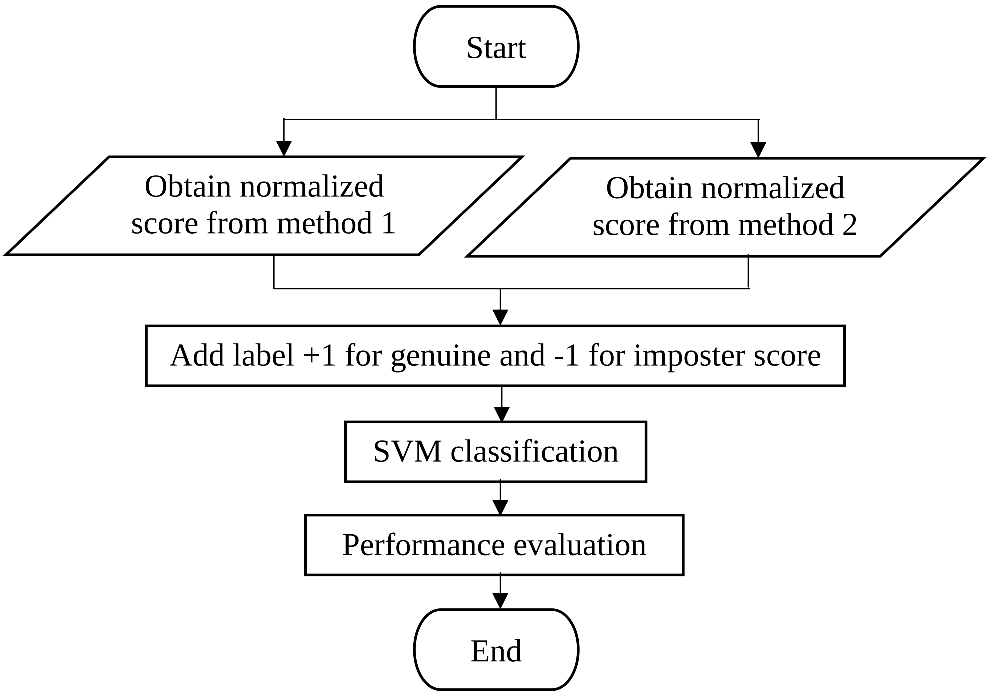


Figure 2.6: Flowchart of score-level fusion using SVM

2.5 Summary

Summaries of feature extraction methods by researchers are listed in Table 2.1. The weaknesses for finger vein feature extraction methods discussed is listed.

Table 2.1: Summaries of feature extraction methods on biometrics

Method	Weakness
Repeated line tracking (Miura et al., 2004)	Predefined probability is required to choose the direction for tracking point to move. The initial tracking point is selected randomly, which might miss some part of the finger vein image.
Principal Component Analysis (PCA) (Wu & Liu, 2011; Yang & Zhang, 2012)	This method proposed may ignore some local detailed information as the PCA extracted global features and directly apply on image classifier.
Linear Discriminant Analysis (LDA) (Wu & Liu, 2011)	This method needs to transform image matrix into high dimension vector as PCA.
Band-Limited Phase-Only Correlation (BLPOC) (Ito et al., 2004; Mohd Asaari et al., 2014)	Unwanted frequency components are included in Phase-Only Correlation (POC). BLPOC is introduced to filter those frequency components. BLPOC only takes the phase into account, neglecting the magnitude of the vein images.

CHAPTER 3

METHODOLOGY

3.1 Introduction

In finger vein verification, the stages involved are image acquisition, pre-processing, feature extraction and matching test. In this project, the main focus is on feature extraction. Spatial pyramid pooling is chosen for learned feature extraction method and BLPOC is used as the method for hand-crafted feature extraction. After the features are extracted, score-level fusion is carried out by using SVM. In this chapter, the procedures for each method are explained in detail.

3.2 Experimental Setup

3.2.1 Database

Finger vein images are required in this project. Thus database of finger vein images used is FV-USM (Rosdi, 2015). There are two sessions for finger vein image collection, each session has 123 different individuals with four different fingers (left index, left middle, right index, and right middle) and each finger has to take 6 times finger vein images. With this, there are total of 5904 images $(2 \times 123 \times 4 \times 6)$.

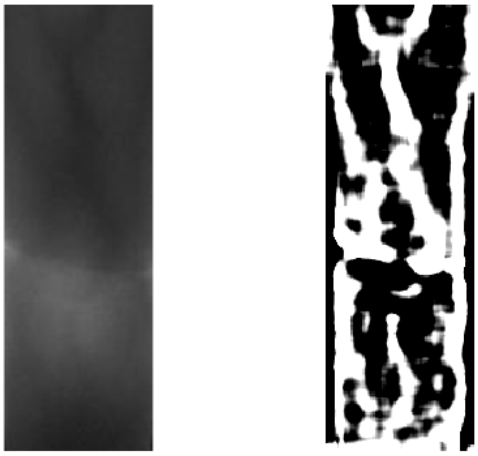
3.2.2 Image Enhancement

Before the feature is extracted, Modified Gaussian Filter (MGF) is applied on finger vein images to improve the performance. The filter's formula is given as (3.1) (Lee, Jung, & Kim, 2011):

$$H(x,y) = b\left(1 - e^{-D^2(x,y)/2D_0{}^2}\right) + c \tag{3.1}$$

where a and b are adjustable variables that can change the magnitude of the filtering mask and $D(x,y)$ is the distance between center of image and a relative position, calculated using formula (3.2). Mask with size of 19×19 pixels is used, with the values of b and c are 13.7379 and -4 respectively. The original finger vein image and enhanced finger vein image are shown in Figure 3.1.

$$D(x,y) = [(x - x_o)^2 + (y - y_o)^2]^{\frac{1}{2}} \tag{3.2}$$



(a) (b)

Figure 3.1: (a) Original finger vein image; (b) Enhanced finger vein image

3.3 Overall Project Flow

The overall flowchart for this project is shown in Figure 3.2. First, the spatial pyramid pooling proposed by Shen et al. (2014) is used for learned feature extraction. Then, the parameter used in the method is modified to suit the application for finger vein verification. After that, hand-crafted feature extraction (BLPOC) is studied and applied, followed by score-level fusion using SVM. Lastly, the performances of all the three methods are evaluated by obtaining frequency distribution graph and Equal Error Rate (EER).

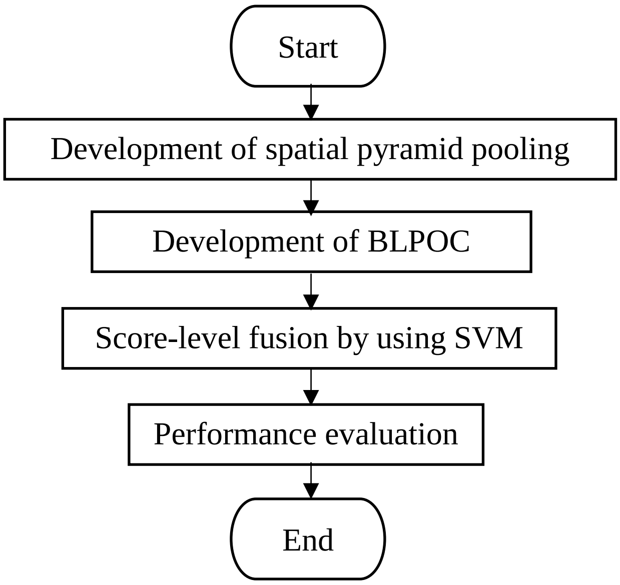


Figure 3.2: Overall project flow

3.4 Spatial Pyramid Pooling

Spatial pyramid pooling is used for feature extraction to process the image by patches instead of global image. The method proposed by Shen et al. (2014) for face recognition, is used in this project because the error is relatively low. The procedure is divided into three different stages, which are pre-processing, spatial pyramid pooling and linear multi-class classification. The method proposed in this project is slightly different from the method proposed by Shen et al. (2014). The process flowchart is shown in Figure 3.3.

In pre-processing stage, suppose that a finger vein image has dimension of 100×300 pixels. First, the image is resized by multiple of 0.2, which produce image with dimension of 20×60 pixels to reduce processing time. Then, the size of local patches is fixed as $A \times B$ pixels to process the image by patch, overlapping each other with a step of 1 pixel. Thus, the image is divided into $l_1 \times l_2$ patches, where $l = [\frac{\text{image dimension} - \text{patch dimension}}{\text{step}} + 1]$, which is also equal to 17×56 patches. Each of the

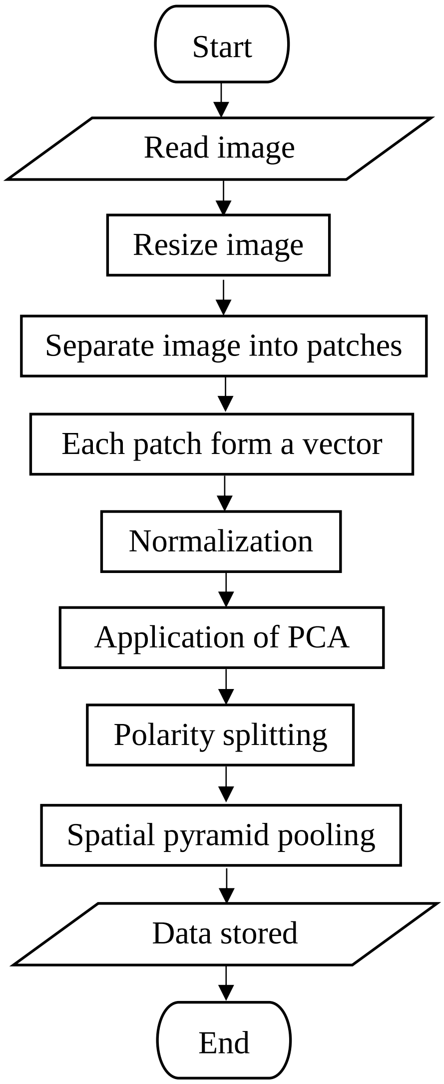


Figure 3.3: Flowchart of algorithm for spatial pyramid pooling method

patches will form a row vector $\boldsymbol{x}$ and normalized using $\widehat{x_i} = (x_i - a)/v$ to increase local brightness and contrast normalization, where x_i is the ith element of x, a and v are the mean and standard deviation of elements of $\boldsymbol{x}$. The rows of vector are then arranged into column. After that, PCA is applied to de-noise and reduce the dimensionality of the final pooled features. The process is then followed by polarity splitting to improve the classification performance. After that, each column will form one patch, by dividing the column into parts and concatenated by the side and prepared for spatial pyramid pooling.

The process of local patch extraction and pre-processing stage is illustrated in Figure 3.4. The preparation for spatial pyramid pooling is illustrated in Figure 3.5.

After local patch extraction and pre-processing, the process is followed by spatial pyramid pooling. In this project, max pooling is used for spatial pyramid pooling as the result obtained from Shen et al. (2014) shows that max pooling is slightly better than average pooling, with the equation given in Chapter 2 (2.13). The feature of finger vein images is pooled layer by layer. First, the maximum value for the patch is pooled. Then the patch is divided into four sub-patches and the maximum value for each sub-patch is pooled. The process is continued according to the pyramid level, for example, in Shen et al. (2014), the 8-level pyramid {1,2,4,6,8,1,12,15} is used. After that, all the maximum values obtained from each patch is arranged in a column. The process for spatial pyramid pooling is illustrated in Figure 3.6. The MATLAB code for spatial pyramid pooling can be view in Appendix A.1.

After the features are pooled, there are two ways used to analyse the result, which are Euclidean distance and indicator. Euclidean distances between features extracted from different images are calculated using (3.3) and used as the score to check Equal Error Rate (EER) later.

$$D = \sum_{i=1}^{I} [(x_{i1} - x_{i2})^2]^{\frac{1}{2}} \tag{3.3}$$

where D is the Euclidean distance, x_{i1} is the ith value pooled from image 1, x_{i2} is the ith value pooled from image 2 and I is the total number of value pooled from images. Next, the indicator is determined. A threshold is set, if $|x_{i1} - x_{i2}|$ is higher than the threshold, it will be indicated as 1. After that, the sum of indicator is calculated, if the value is high, most probably the image is imposter and low for genuine. In this project, the threshold set is 0.006. The results obtained from both Euclidean distance and indicator are analysed, the score using indicator shows lower error. Thus, indicator is used in this project.

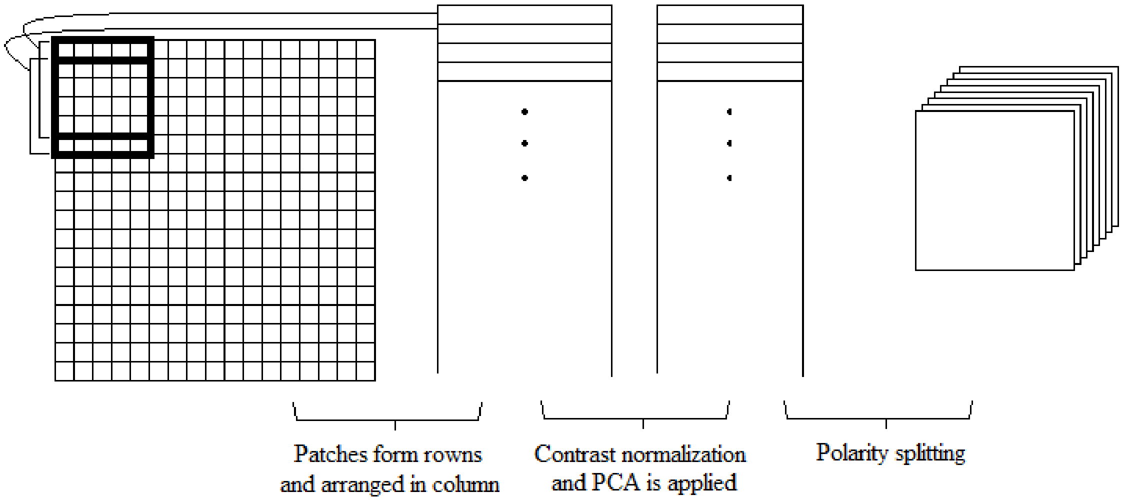


Figure 3.4: Process flow for local patch extraction and pre-processing (Shen et al., 2014)

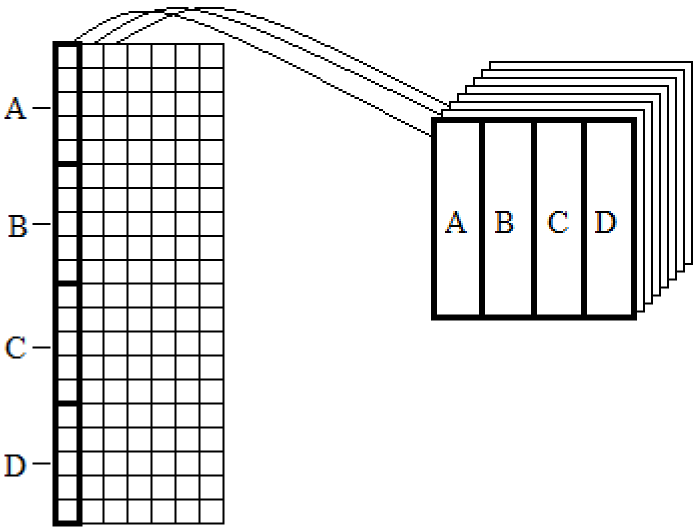


Figure 3.5: Preparation for spatial pyramid pooling