International Journal of Engineering Trends and Technology- Volume4Issue2- 2013

Fingerprint Based Gender Classification Using 2D Discrete Wavelet Transforms and Principal Component Analysis

Rijo Jackson Tom[1] T.Arulkumaran [2]

[1.] M.E Scholar
Sathyabama University, Chennai-119, India

[2.] Assistant Professor
Sathyabama University, Chennai-119, India

***Abstract*— Fingerprint evidence is undoubtedly the most reliable and acceptable evidence till date in the court of law. Fingerprints are obtained at the site of crime and in many old monuments and in excavated things. Estimating the gender of fingerprints is an emerging field and many methods using the fingerprint physical features like the ridge count and the ridge thickness have been used so far. Due to the immense potential of fingerprints as an effective method of identification an attempt has been made in the present work to analyze their correlation with gender of an individual using frequency domain technique and a pattern recognition technique. The combined processing has provided better results. This paper aims in using 2D- Discrete Wavelet Transform (DWT) and Principal Component Analysis (PCA) combined to classify gender using an obtained fingerprint. The minimum distance method was used as a classifier. Fingerprints of 200 males and 200 females belonging to the various age groups were taken for analysis. The experimental results show good for trained database. It was found that increasing the database population in each category improves the performance of the system.**

***Keywords* – DWT, PCA**

I. INTRODUCTION

FINGERPRINTS are one of the most mature biometric technologies and are considered legitimate proofs of evidence in courts of law all over the world. Based on the varieties of the information available from the fingerprint we are able to process its identity along with gender, age and ethnicity [1]. Within today's environment of increased importance of security and organization, identification and authentication methods have developed into a key technology. Such requirement for reliable personal identification in computerized access control has resulted in the increased interest in biometrics [6].

A Fingerprint is the representation of the epidermis of a finger; it consists of a pattern of interleaved ridges and valleys. Fingertip ridges evolved over the years to allow humans to grasp and grip objects. Like everything in the human body, fingerprint ridges form through a combination of genetic and environmental factors. This is the reason why even the fingerprint of identical twins is different (Maltoni and Cappelli, 2006). The concept of fingerprint pattern being studied has been of significant use over time, when scanning it involves the conversion of fingerprint by small portion of light solid-state devices into alphanumeric formula (Galton, 1982).

Fingerprint identification algorithms are well established and are being implemented all over the world for security and person identity. Very few attempts have been made to classify the gender from an obtained fingerprint. This is helpful for anthropologists for classifying gender from the fingerprints they obtain from excavated articles and for crime investigators for minimizing the rage of the suspects.

The gender of the person can be judged using the fingerprint of that concern person based upon the count of the ridges of the fingerprint. The average ridge count is slightly higher in males than in females, with high standard deviation among subjects of both genders. Epidermal ridges and their arrangement (dermatoglyphic patterns) exhibit a number of properties that reflect the biology of an individual. Dermatoglyphic features statistically differ between the sexes, ethnic groups and age categories [4].

The paper is aimed in developing an algorithm for classifying the gender through fingerprint obtained.

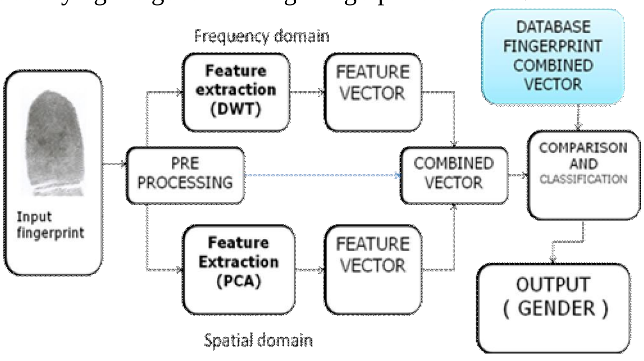


Fig. 1. Generalized Block Diagram

Fingerprints are acquired real time and from the database sources. A fingerprint gender identification system constitutes of digital images of fingerprint as its input which is then transformed. The algorithm can be developed using MATLAB programming language. We use 2D- Discrete Wavelet Transform (DWT) and Principal Component Analysis (PCA) combined to classify gender using his/her fingerprint.

The obtained fingerprint goes through preprocessing stage for enhancement and removing the noise. After preprocessing the fingerprint goes through two levels of feature extraction, one is frequency domain feature vector obtaining by undergoing Wavelet decomposition and second is by spatial level undergoing PCA. The next step is to combine the two vectors and this is compared with the database for the minimum distance and classifies the fingerprint as to which class it lies.

The rest of the paper is organized as follows: literature survey is given in section II. Proposed system is discussed in section III. Gender classification in section IV.

II. Literature Survey

A. Overview of fingerprinting

Human fingerprints have been discovered on a large number of archeological artifacts and histological items. Although these findings provide evidence to show that ancient people were aware of the individuality of fingerprints, it was not until the late sixteenth century that the modern scientific fingerprint technique was first initiated (Jain, et al, 2003). In 1686, Marcello Malpighi, a professor of anatomy at the University of Bologna noted in his writings the presence of ridges, spirals and loops in fingerprints.

Henry Fauld, in 1880, was the first to scientifically suggest the individuality of fingerprints based on an empirical observation. In the late nineteenth century, Sir Francis Galton conducted an extensive study on fingerprints; he introduced the minutiae features for fingerprint matching in 1888. An important advance in fingerprint recognition was made in 1899 by Edward Henry, who established the well-known Henry system of fingerprint classification (Lee and Gaensslen, 2001).

In the early twentieth century, fingerprint recognition was formally accepted as a valid personal identification method and became a standard routine in forensics (Lee and Gaensslen, 2001). Fingerprint identification agencies were set up world-wide and criminal fingerprint databases were established (Lee and Gaensslen, 2001). Various fingerprint recognition techniques, including latent fingerprint acquisition, fingerprint classification and fingerprint matching were developed. For example, the FBI fingerprint identification division was set up, in 1924, with a database of 810,000 fingerprint cards (Federal Bureau of Investigation, 1984).

In 1890, Alphonse Bertillon studied body mechanics and measurements to help in identifying criminals. The police used his method, the Bertillon age method, until it falsely identified some subjects. The Bertillon age method was quickly abandoned in favor of fingerprinting, brought back into use by Richard Edward Henry of Scotland Yard. Karl Pearson, an applied mathematician studied biometric research early in the 20^{th} century at University College of London. He made important discoveries in the field of biometrics through studying statistical history and correlation, which he applied to animal evolution. His historical work included the method of moments, the Pearson system of curves, correlation and the chi-squared test.

The system was first used in India in 1858 by Sir William Herschel to prevent impersonation, but the credit is given to Sir Francis Galton for having it systematized for the identification of criminals. His system was officially adopted in England in 1894, and was further modified by Sir Edward Henry. Afterwards the studies have been conducted on fingerprint ridges mainly its types, classification, methods of lifting fingerprints, recording of fingerprints and materials used to develop fingerprint (Gungadin 2007).

B. Recent researches in Fingerprint Classification

Fingerprint identification and classification has been extensively researched in times past, however very few researchers have studied the fingerprint gender classification problem, (Acree 1999) used the ridge density, which he defined as the number of ridges in a certain space; it was shown that the females have higher ridge density (Acree1999). (Kralik 2003) also showed that the males have higher ridge breadth,) which was defined as the distance between the centers of two adjacent valleys), than females.

Two studies showed that the males have higher ridge count than the females (Hall, Timura 2003) and (Cote, Earls, Lalumiere 2002). It was shown that both males and females have higher rightward directional asymmetry in the ridge count, with the asymmetry being higher in males than females (Hall and Kimura 2003), and higher incidence of leftward asymmetry in females. (Cote et al 2002). Female's fingerprints are significantly of lower quality than male fingerprints (Hicklin and Reedy. 2002). While the appearance of white lines and scars in fingerprint images are very common in housewives (Wendt 1955).

III. Proposed System

The following paper aims in creating a system that is used to classify gender from an obtained fingerprint. Fingerprints are acquired real time and from the database sources. A fingerprint gender identification system constitutes of digital images of fingerprint as its input which is then transformed. We use 2D- Discrete Wavelet Transform (DWT) and Principal Component Analysis (PCA) combined to classify gender using his/her fingerprint.

The obtained fingerprint goes through preprocessing stage for enhancement and removing the noise. After preprocessing the fingerprint goes through two levels of feature extraction, one is frequency domain feature vector obtaining by undergoing Wavelet decomposition and second is by spatial level undergoing PCA. The next step is to combine the two vectors and this is compared with the database for the minimum distance and classifies the fingerprint as to which class it lies.

The database for our paper contains 547 individual fingerprints of male and female of different ages between 12-64. The database fingerprints goes through the process and the feature vectors are stored in the database and is used for classification.

A. Preprocessing

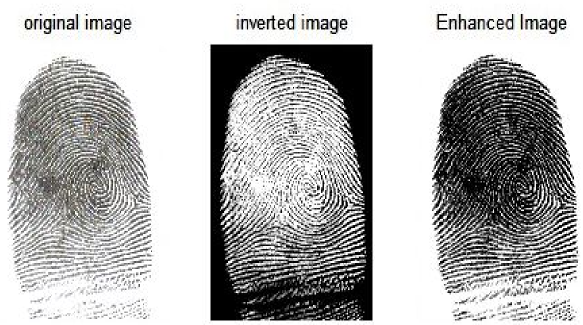
Image enhancement processes consist of a collection of techniques that seek to improve the visual appearance of an image or to convert the image to a form better suited for analysis by a human or a machine. Enhancements techniques like contrast enhancement, histogram equalization, biniarization, thinning and inverting are used as per the requirement of the image to be enhanced. The fingerprint is resized to 512x512. The fingerprint image obtained undergoes image enhancement for improving quality of the ridges and valleys. The input image which is grayscale is converted into binary. The output of preprocessing is shown in Fig 2. After preprocessing the fingerprints undergo further processing. Enhancement methods changes from fingerprint to fingerprint and for different databases. Poor quality fingerprints obtained can be enhanced for the betterment of the algorithms.

B. DWT Feature Vector generation:

The fingerprint image undergoes discrete wavelet transformation for obtaining the feature vector. Wavelets have been used frequently in image processing and used for feature extraction, denoising, compression, face recognition, and image super-resolution. Two dimensional DWT decomposes an image into sub-bands that are localized in frequency and orientation. The decomposition of images into different frequency ranges permits the isolation of the frequency components introduced by “intrinsic deformations” or “extrinsic factors” into certain subbands. This process results in isolating small changes in an image mainly in high frequency subband images. Hence, discrete wavelet transform (DWT) is a suitable tool to be used for designing a classification system.

The obtained image is decomposed using the analysis filter bank and the low frequency and the high frequency bands are separated as shown in the Fig. 3.

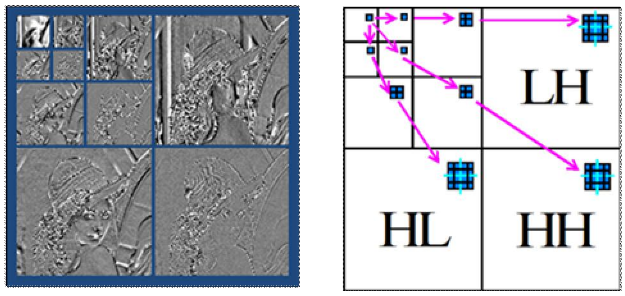
The 2-D wavelet decomposition of an image is results in four decomposed sub-band images referred to as low–low (LL), low–high (LH), high–low (HL), and high–high (HH). Each of these sub-bands represents different image properties. Typically, most of the energy in images is in the low



frequencies and hence decomposition is generally repeated on the LL sub band only (dyadic decomposition).

Fig. 2 Enhanced Fingerprints
Fig 3 Wavelet Decomposition

For k level DWT, there are (3*k) + 1 sub-bands available. The energy of all the sub-band coefficients is used as feature vectors individually which is called as sub-band energy vector (E_k). The energy of each sub-band is calculated by using the equation



$$E_k = \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}|X_k(i,j)| \qquad (1)$$

Each fingerprint undergoes six level decomposition after preprocessing as shown in Fig. 3. At each level we get subbands and their energy is calculated. We get a total of 19 subbands and the energy vector E_k will be 1x19 vector at the end of the six level decomposition for each fingerprint. All fingerprints in the database undergo the decomposition and the energy vector of all the images is stored.

C. PCA Feature Vector generation:

The preprocessed fingerprint also undergoes PCA algorithm for obtaining the feature vector in spatial domain. PCA algorithms are generally implemented for pattern recognition systems. Principal component analysis involves a mathematical procedure that transforms a number of possibly correlated variables into a smaller number of uncorrelated variables called principal components. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible. It is also named as the Karhunen–Loeve transform which is also called as (KLT) the Hotelling transform.

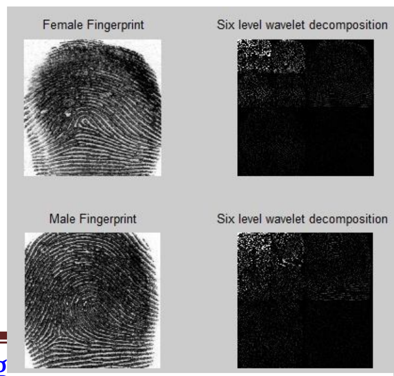


Fig. 4 Six level decomposition of fingerprint

PCA involves the calculation of the eigen value decomposition of a data covariance matrixhttp://en.wikipedia.org/wiki/Covariance_matrix or singular value decomposition of a data matrix, usually after mean centering the data for each attribute. The results of a PCA are usually discussed in terms of component scores and loadings. PCA involves the calculation of the Eigen value decomposition of a data covariance matrixhttp://en.wikipedia.org/wiki/Covariance_matrix or singular value decomposition of a data matrix, usually after mean centering the data for each attribute. The results of a PCA are usually discussed in terms of component scores and loadings.

An eigen vector of a given linear transformation is a vector which is multiplied by a constant called the eigen value as a result of that transformation. The direction of the eigenvector is either unchanged by that transformation (for positive eigen values) or reversed (for negative eigen values).

Every fingerprint in the database undergoes the PCA for obtaining the eigenvector. The eigenvector is 512x 512 size and is stored as another feature vector of the fingerprint.

D. Combined Feature Vector:

After undergoing DWT and PCA the feature vectors are stored separately. The next step is to combine both the feature vector into a single vector which stores the frequency domain and spatial domain information of a fingerprint.

The 19 feature vector from the DWT and the 512 feature vector from PCA are combined to form 1 x 531 feature vector for a single fingerprint. Similarly for all the fingerprints in the database this procedure is followed and a database feature vector is created which contains all the feature vector of the images in the database. If there are *n* fingerprints in the database then the size of the database feature vector will be *n* x 531.

A graph was plotted after obtaining the combined feature vector; it is shown in Fig 5. The graph shows a demarcation between Male and Female prints meaning to say that there are differences between male and female fingerprints even when analyzed using frequency domain analysis. Till date spatial features where used for analysis and is well established. We propose this system for the use of frequency domain analysis for the purpose.

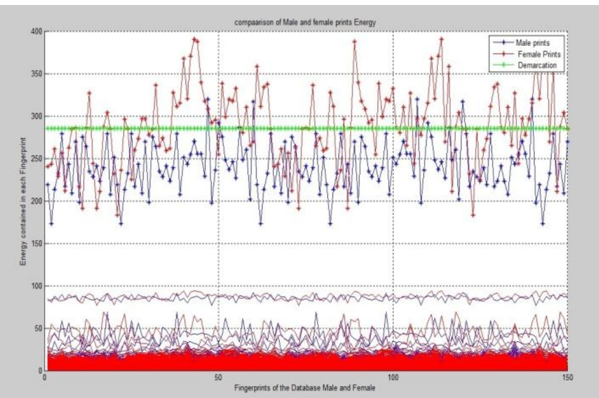


Fig. 5 Database feature vector of Male and female

IV. GENDER CLASSIFICATION

We have used 200 male fingerprints and 200 female fingerprints downloaded from the database of Department of Biometric research, University of Ilorin. All these fingerprints are obtained by means of optical scanner. For gender classification the database contains only two classes Male and Female. The fingerprints of both classes are grouped and kept as the database fingerprints.

Once the desired database is formed all the fingerprints in it undergoes the feature vector extraction as explained in the previous section.

Steps to be followed for Gender classification using the query fingerprint:

1. The fingerprint undergoes pre processing and is resized to 512x512.
2. The fingerprint undergoes Wavelet Decomposition and the 19 feature vector is obtained.
3. The fingerprint from the pre processing stage also undergoes PCA Eigen vector feature extraction. This provides the eigen vector of 512 size.
4. Now the features vectors are combined in total 19+512 = 531 vectors are obtained from the fingerprint.
5. This fingerprint feature vector is classified using the minimum distance classifier and this undergoes the gender classification.

A. Gender Classification:

Query fingerprint whose gender has to be classified also undergoes the feature extraction; and 1 x 531 feature vector is obtained. This query feature vector is compared with all the database feature vector of both the classes. Here for our approach we have used a simple classifier Minimum Distance Classifier which uses the Euclidian distance measure for finding the distance between the query and the database.

Once the distance is found between the query and the database, classification comes into picture. We analyze to see as to which class does the following query fingerprint lies closest.

The gender classification algorithm was implemented and the output was tested.

Fig. 6 Gender Classification

The classification of the gender is done using the minimum distance classifier. The feature vector of the male and the female fingerprints are stored separately. The query feature vector obtained from the above process is compared with both the classes. The classification algorithm developed here is simple. By analyzing the minimum distance in the class the fingerprint is classified accordingly.

V. PERFORMANCE ANALYSIS

The algorithm was written in MATLAB 7.5 and was run on Intel Core 2 duo processor 1.66 GHz with 1 GB memory. The process of feature vector extraction for the database is time consuming. We have used ¾th of the database fingerprints for training and the rest was used for testing purpose. The result of the gender classification is shown in Fig 7.

Results show us that frequency domain analysis of fingerprints can also be followed in future for the study of dermatoglyphic.

The overall success rate in gender classification is around 70% for our database.

The algorithm was implemented only on optical scanned prints. With optical scanned prints in database and ink prints as query the algorithm was not much efficient because of the lack of training.

It has been found that changing the database also improves the classification process. The number of database prints is an important criterion. Experimental results show an overlap between the male and the female fingerprints in some cases. There are changes that fingerprints of females can be similar to male prints.

The fingerprints of different age groups vary in size and patterns and thickness of ridges and valleys. The fingerprints of people from various ethnic groups vary. An algorithm for compressing the huge database of fingerprints has to be developed and the database of the feature vectors have to be coded to provide a simpler database structure to reduce the complexity in calculations.

VI. FUTURE WORKS

The algorithm that was developed was simple which was to study how well frequency domain analysis help in the classification. In future the algorithm has to be improved with a good classification algorithm like neural networks. It has been found that improving the database is an important criterion for good classification and estimation. Ink prints, optical scanned prints and prints from artifacts all can be used in the database. In future fingerprints from different ethnic groups have to be collected for a large scale study.

An algorithm has to be developed for the compression of the fingerprints and its feature vectors. Databases from different population groups have to be collected for storing the feature vector.

Classification algorithms can be developed that can be provided with the fingerprint feature vector for training. This will improve the strength of the algorithm.

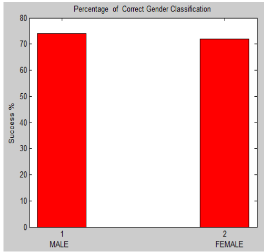


Fig. 7 Success rate of Gender Classification

REFERENCES

[1] P. Gnanasivam & Dr. S. Muttan " Estimation of Age Through Fingerprints Using Wavelet Transform and Singular Value Decomposition" *International Journal of Biometrics and Bioinformatics (IJBB), Volume (6) : Issue (2) :* pp 58-67. 2012

[2] Ritu Kaur and Susmita Ghosh Mazumdar "Fingerprint based gender identification using frequency domain analysis" *International Journal of Advances in Engineering & Technology, Vol. 3, Issue 1,* pp. 295-299, March 2012

[3] Pankanti, S. Prabhakar, and A. K. Jain, "On the individuality of fingerprints," *IEEE Trans. Pattern Anal. Mach. Intell.",* vol. 24, no. 8, pp. 1010–1025, Aug. 2002.

[4] E.O. Omidiora, O. Ojo, N.A. Yekini, T.O. Tubi "Analysis, design and implementation of human fingerprint patterns system "towards age & gender determination, ridge thickness to valley thickness ratio (RTVTR) & ridge count on gender detection" *International Journal of Advanced Research in Artificial Intelligence, Vol. 1, No. 2,* pp 57- 63, 2012

[5] http://www.unilorin.edu.ng/step-b/biometrics

[6] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, "Handbook of Fingerprint Recognition", first ed., *Springer*, New York, 2003.

[7] Miroslav Kralik, Vladimir Novotny, "Epidermal Ridge Breadth: An Indicator of Age and Sex in Paleodermatoglyphics", *Variability and Evolution*, Vol. 11, pp. 5-30, 2003.

[8] K. Thaiyalnayaki, S. A. Karim, P. V. Parmar, "Finger print Recogntion Using Discrete Wavelet Transform", *International Journal of Computer Applications*,pp.96-100,Vol.1,No.24, 2010

[9] A. K. Jain, S. Prabhakar and L. Hong, "A multichannel approach to fingerprint classification," *IEEE Transactions on PAMI*, pp. 48359, Vol.21, No.4, Apr.1999.

[10] Manish Verma and Suneeta Agarwal, " Fingerprint Based Male-Female Classification.'' in Proceedings of the *International Workshop On Computational Intelligence in Security for Information Systems* (CISIS'08), Genoa, Italy, 2008, pp.251-257.

[11] D.Maio and D.Maltoni, "Ridge-line density estimation in digital images," in Proceedings of the *14th International Conference on Pattern Recognition (*ICPR), 1998, pp. 534– 538.