UTM Computing Proceedings
Innovations in Computing Technology and Applications
Volume: 1 | Year: 2016 | ISBN: 978-967-0194-81-3

Bio Digital ID: Fingerprint Feature Extraction Using Invariant Moment

Mohamed Azreen Mohamed Zaini[1], Anazida Zainal*[2]

Department of Computer Science, Faculty of Computing,
Universiti Teknologi Malaysia,
81310 Johor Bahru, Johor, Malaysia
[1]ayinbzaini@gmail.com, [2]anazida@utm.my

Abstract

The advancement of technology has strong influence in spreading information and data across the globe. Wide variety of data and information such as image file, video, and audio file goes into our lives and out every day. As digital multimedia works (video, audio and images) become globally available for retransmission, reproduction and publishing over the internet, it contributes to the increment of legal copyright issues, for instance unauthorized copy and distribution. Digital watermarking is one of the way to fight against this violation. In digital watermarking, a data or message known as watermark or signature is embedded into a multimedia object (image, audio or video) which later can be extracted for ownership verification or authentication. Recent study shows that a data or message is expected to have the characteristic of imperceptibility or cannot be known to others and contains the author's information. Therefore, this research aims to study the combination of biometric system which is the use of human fingerprint features extracted using Invariant Moment and hashing algorithm. The generated hash value later can be used as data or message in digital watermarking. This study includes the analysis of current approaches or techniques in fingerprint biometric system, implementation of invariant moment for feature extraction and existing hashing method. Data used in this research are divided into two namely for validation (FVC2004) and testing (20 SCSR Student). The result shows that hash value created from 20 average value of invariant moment has no duplication. Thus proved the proposed method is suitable to be used as data or message in digital watermarking.

Keywords: Invariant Moment, Absolute Distance, Hashing, Feature Extraction

1.0 Introduction

Nowadays, due to technology rapid growth, digital product have been very popular among us. Many platform such as Amazon, Spotify, and ITunes are some the ways of gaining this product legally. However, as digital multimedia works (video, audio and images) become available for retransmission, reproduction, and publishing over the Internet, a real need for protection against unauthorized copy and distribution is increased. As the result, production teams, video production teams, recording labels, and software developers are the people who checks are being cut due to this file sharing.

Bonner and O'Higgins (2009) stated that a few music company such as Sony Music Entertainment, Warner Music Group, and Universal Music Group already taking step to solve the problem by filing lawsuits against illegal consuming individual websites that allow

file sharing, However this action , aren't actually affective because music and move sales are steady decreasing, while illegal downloading steady increases. Thus, it is important to provide others copyright protection against these violation. One of the ways is to make a digital watermarking into the videos, images and audios product.

According to Chakradhar (2012), digital watermarking is the practise of embedding data or message called watermark or signature into a multimedia object (image or audio or video) so that the watermark can be extracted for ownership verification or authentication. This technology is becoming important due to the popularity of usages of audio on web. Adiwijaya et al (2013) stated that, data or message that is used in the watermarking technique cannot be known directly (imperceptible). He also stated that the data can be an author's information or other information that sign a copyright of that work. Thus, both of this requirement need to be follow in order to achieve high copyright protection.

Therefore, to fulfill the requirement, combination of biometric system which is human fingerprint feature that extracted using Invariant Moment and hashing algorithm seem to be the best method. The feature vector that extracted using invariant moment is hashed using hashing algorithm in order to create a hash value that can be used as data or message in digital watermarking.

Several objectives are identified to achieve the project which are: (i) to extract fingerprint feature using Invariant Moment, (ii) to develop modified MD5 algorithm that used as hashing algorithm to create Bio Digital Id based on average value of invariant moment, and (iii) to evaluate the matching accuracy based on Equal Error Rate (EER).

2.0 Methodology

The process includes four phases such as Image Pre-processing, Feature Extraction and Enrolment, Matching and Hashing and Performance Measurement as shown in Figure 1. In phase 1 , image pre-processing consist of 4 stage such as Input fingerprint image, improve clarity of the ridge structure in the fingerprint image using STFT analysis, detection of reference point using complex method and improve the accuracy in invariant moment feature using region of interest (ROI) . In phase 2, based on ROI, five fingerprint image will be extracted using invariant moment analysis and calculate the average value of the feature vector as baseline for the matching process. In phase 3, feature vector produced will be matching with feature vector stored in the database using absolute distance and the feature vector will be used in process creating Bio digital id by hashing using modified algorithm. In phase 4, the matching accuracy will be evaluated based on the False Acceptance Rate (FAR) and False Reject Rate (FRR).

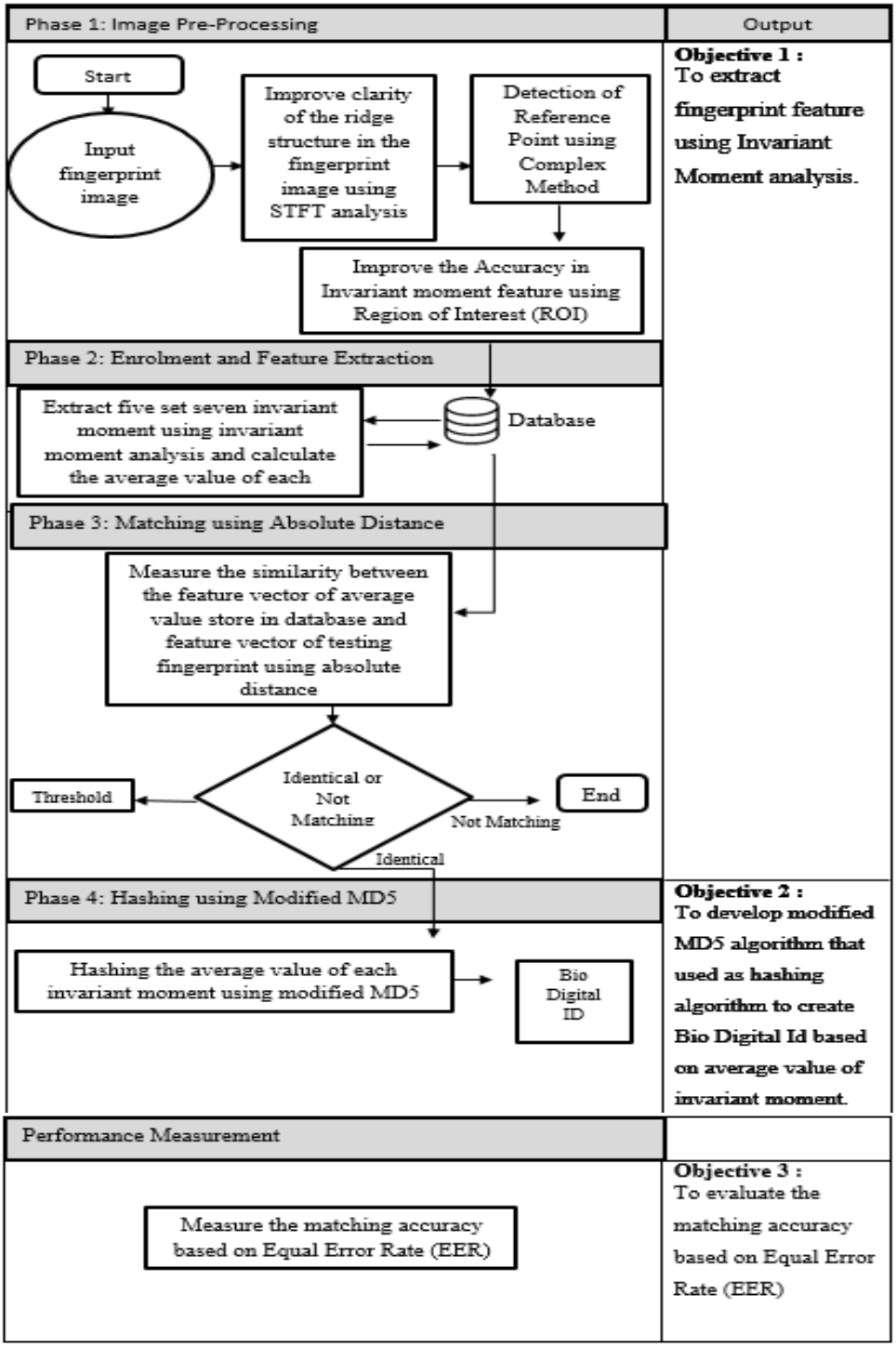

Figure 1 Research Framework

(a) Image Pre-processing

The phase will include 4 processes such as Input fingerprint image, improve clarity of the ridge structure in the fingerprint image using STFT analysis, detection of reference point using Complex Method and improve the accuracy in invariant moment feature using Region of Interest (ROI). The output of this phase is image of size 64 x 64 ROI with reference point at center and the output will be use next in phase 2 (Feature Extraction and enrolment).

(i) Input fingerprint image

The fingerprint acquisition is done using an optical fingerprint sensor ZK4500 which captures the fingerprint image with a resolution of 512 dpi and a 340 x 340 pixel grayscale size. A database of 114 fingerprint images was created corresponding to 19 different people, that is, 6 images of each person.

(ii) Improve clarity of the ridge structure in the fingerprint image using STFT analysis.

After the process to acquire fingerprint image, the STFT enhancement will used to enhance the fingerprint image upon the quality of the input fingerprint image. In this process, the algorithm for the image enhancement consists of two steps such as STFT analysis and enhancement. Figure 2 and Figure 3 show the fingerprint image before and after the enhancement.

Figure 2 Before enhancement

Figure 3 After enhancement

(iii) Detection of Reference Point using Complex Method

After the fingerprint image has been enhanced using STFT enhancement, the reference point will be determined on the enhanced fingerprint image-using complex filters method. In this the research, the reference point will be defined as point that have maximum curvature on the convex ridge (Liu, 2005). The reference point also always located in the central area of fingerprint known as core shown in Figure 4.

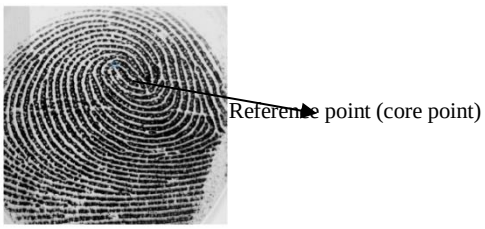

Figure 4 Reference point

(iv) Improve the Accuracy in Invariant Moment Feature using Region of Interest (ROI).

Fingerprint image will be segment by cropping the image based on the reference point determined in the previous process. Instead use entire fingerprint image, this research use certain area around the reference point (or ROI). The purpose of this process is to acquire accurate invariant moment features in the next process based on determination of ROI. In the research, the fingerprint image will be cropped with a size of 64x64 ROI with a reference point at the center.

Figure 5 Region of Interest

(b) Feature Extraction and Enrolment

Five set of seven invariant moments will be extracted from ROI and calculate the average value of the five set. The average value is used as baseline for the matching process later. In this research Hu Invariant Moment (1962) is used process. The result is a set of absolute orthogonal moment invariants that can be used for scale, position, and rotation invariant pattern identification.

$$
\begin{aligned}
\phi_1 &= \eta_{20} + \eta_{02} \\
\phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
\phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - 3\eta_{03})^2 \\
\phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
\phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
\phi_6 &= (\eta_{20} - \eta_{02})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
&\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})
\end{aligned}
\qquad (1)
$$

$$
\begin{aligned}
\phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&\quad + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]
\end{aligned}
$$

The enrolment process will store the information in the database such as respondent name, matrix number and average value of five set of seven invariant moment. The database will be use next in matching process. If the respondent already exists in database, the process will continue to the next phase.

(c) Matching using absolute distance

Absolute distance will be used in to measure the similarity between the feature vectors of the two fingerprints to be matched such testing fingerprint and average value stored in the database. According to Yang and Park (2008) below are the steps to conduct matching proses:

Step 1: Define average value of seven invariant moment in database as M = $M_1 M_2 \ldots . M_7$ and seven invariant moment of testing fingerprint as N = $N_1 N_2 \ldots . N_7$

Step 2: Find the difference vector V_d

$$
V_d = \left(\frac{|a_1 - b_1|}{\max(a_1, b_1)}, \frac{|a_2 - b_2|}{\max(a_2, b_2)} \ldots\ldots \frac{|a_n - b_n|}{\max(a_n, b_n)} \right)
$$

Step 3: Define the absolute distance

$$R_m = \sum_{i=1}^{n} \frac{|a_i - b_i|}{\max(a_i, b_i)}$$

[[(3)

Step 4 : Determine the matching result.

Accept if R_(m) < T_m

Reject if R_(m) > T_m

In this research, T_m will be set to 20

(d) Hashing using modified md5 algorithm

The average value of the five set feature vector store in database will used as input to create bio digital id by hashing using modified md5 based on MD5 architecture and combination mathematical equation such as standard deviation and modulus operation. The processing consists of the following steps:

Step 1: Create 4 word buffer

Convert the Moment Invariant 1 (M1) to the hexadecimal and divide by 4 part equally from left to right to create four-word buffer. The Buffer will initialized as A, B, C and D as shown in Figure 6. Each of this variable is a 32 bit number.

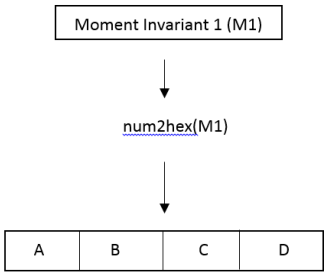

Figure 6 Four Word Buffer

Step 2: Improve Unique Value

Find the standard deviation of the Invariant Moment based on the combination that had been assigned. The purpose of this step is to improve the unique value that will be generated. Then convert the value generated using num2hex to produce hexadecimal number.

$$H_i = \text{num2hex}\left(\left[\sqrt{\frac{\sum(x-\mu)^2}{n-1}}\right]\right) \qquad (4)$$

M2 and M7 = H_1

M3 and M6 = H_2

M4 and M5 = H_3

M1, M2, M3, M4, M5, M6, M7 = H_4

Step 3: Concatenate to produce new hexadecimal value

Concatenate hexadecimal value from step 1 and step 2 to produce new hexadecimal value. Then convert the new hexadecimal using hex2dec to produce decimal value.

D_1 = hex2dec (AH_1)

D_2 = hex2dec (BH_2) (5)

D_3 = hex2dec (CH_3)

D_4 = hex2dec (DH_4)

Step 4: Find modulus to produce integer value

In addition to improve the uniqueness and cryptographic measure to the value generated on step 3, user the modulus operation to produce new integer. Then concatenate the value start from V_1 until V_4. E will become the new value.

V_1 = D_1 mod (power of D_1)

V_2 = D_1 mod (power of D_2) (6)

V_3 = D_1 mod (power of D_3)

V_4 = D_1 mod (power of D_4)

E = $V_1V_2V_3V_4$

The contents of the four buffers (A, B, C and D) are now mixed with Value of E to produce new Hash Value as shown in Figure 7.

D	A	E	B	C

Figure 7 New Hash Value

3.0 Result

In this research, different threshold is tested in order to identify peak threshold that show most accurate result as shown in Table 1. The Equal Error Rate (EER) is used as a performance indicator. The EER indicates the point where the False Rejection Rate (FRR) and False Acceptance Rate (FAR) are equal as shown in diagram 4.2. In this research, the value of threshold is set to 20, which show optimum rate of FRR and FAR.

Table 1: Identifying Peak Threshold

Threshold	FRR (%)	FAR (%)	EER (%)
0.24 (Standard Deviation Calc)	48.3	45	42
1	41.6	40	
10	20	60	
20	16.7	60	
30	15.8	65	
40	12.5	70	
50	10	70	
60	10	75	

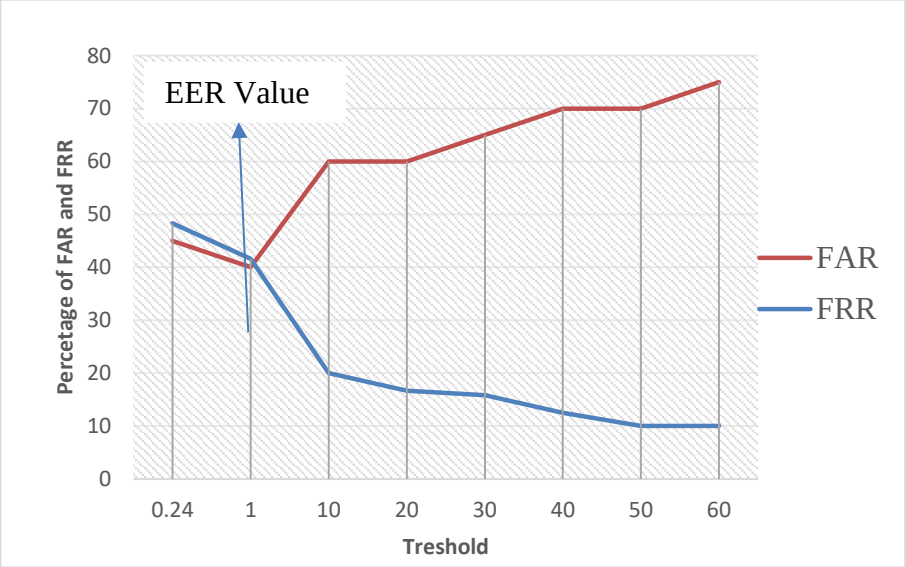

Figure 8 Graph of FRR and FAR against Threshold

The hash value shown in Table 2 indicates that there is no duplication hash value between 20 average values that use as baseline in the matching process. This show the hashing algorithm can be used to produce unique hash value.

Table 2: Hash Value

96663fe311441225dc8258	ff3d3fda226001f37f774
e9c83fdc418118847801dc	b6b23fd7328170c8ee9d4
5bf73fe68342066b94e7c	55dc3fe121201118d29d199f
c5a53fe0916131d42e8458	b73a3fdf6181622625bfeb5
cf903fe68013103735606e	2a5b3fd712181033d850e49
d3463fdc60819e778e8fd	03603fe72221084f6b9cff
a7d63fe48412972c4aabb	3f8a3fe0121013321b02cdf
60493fe099010bafbdec4	30933fe031122136a4c4f5b
bcc23fed101515658b3174c	aea53fdd1384267ce6429
bf133ff219120176883fcf2	3e1a3fdc138159dda6297f

4.0 Discussion

In this research, based on Figure 8 the value of FRR decrease gradually as the threshold increase. This indicate the greater the value threshold, the lower the number of rejected genuine claim. When the threshold approaching value 50 and 60, the graph did not show any changes on the percentage, this is because mostly matching value produced is lower than 60. The value of FRR in this research, show that the matching technique has higher accuracy to detect fingerprint whether identical or not when tested with same person.

In FAR, the value increases gradually as the threshold increase. This result shows that the greater the value threshold, the higher the number of accepted imposter claims. The value of FRR in this research, show that the matching technique has lower accuracy to detect fingerprint whether identical or not when tested with different person.

As overall in performance indicator, the value of ERR is 42%.The accuracy of the matching technique is still unsatisfactory, which is because different ROI of the same fingerprint will generate a much different invariant moment feature vector which can cause a false reject. This due to the manually process on identifying ROI that led to human error. If the ROI is identifying automatically, the matching accuracy would be much higher. Despite low matching accuracy, the hash value shown in Table 2 indicate that there is no duplication hash value between 20 average values that use as baseline in the matching process. This show the hashing algorithm can be used to produce unique hash value that can be use as data or message in digital watermarking.

5.0 Conclusion

In this paper, texture image extraction technique which is Invariant Moment is used to extract fingerprint feature vector. The experimental result shows that the absolute distance can use to discriminate different fingerprints. The matching accuracy technique is still unsatisfactory to the manually process on identifying ROI that led to human error. If the ROI is identifying automatically, the matching accuracy would be much higher. The hashing algorithm that been used proved that hash value generated show uniqueness. This show the hashing algorithm can be used to produce unique hash value that can be use as data or message in digital watermarking.

References

Yang, J.C., Shin, J. W., Park, D. S. (2007). "Fingerprint Matching Using Invariant Moment Features", Lecture Notes in Artificial Intelligence(LNAI 4456), Springer, Berli.

Jain A. K., Uludag,U. (2004). Attacks on biometric systems: a case study in fingerprints, Proc. SPIE-EI 2004, Security, Seganography and Watermarking of Multimedia Contents VI, San Jose, CA, pp.622–633

Leon, J., Sanchez, G., Aguilar, G., Toscano, L., Perez, H., Ramirez, J.M. (2009). Fingerprint verification applying invariant moments, The IEEE International Midwest Symposium on Circuits and Systems, pp. 751-757

Yang,J. C. and Park, D. S.(2008)."Fingerprint verification based on invariantmoment features and nonlinear BPNN", Int. J. ControlAutomSyst.,vol.6,no.6,pp.800-80

Adiwijaya, A., Novraditya, E., Baihaqi, F.N.,Wisesty, U.N. (2013) "Copyright Protection in Audio File Using Watermarking Approach Based on Wavelet-SVD", Applied Mechanics and Materials, Vols. 321-324, pp. 1191-1195.

Bonner, S., & Eleanor O'Higgins. (2010). Music piracy: Ethical perspectives. Management Decision, 48(9), 1341-1354.

Chaudhari, A. S., Patil, S. S. (2013). A Study and Review on Fingerprint Image Enhancement and Minutiae Extraction", IOSR Journal of Computer Engineering (IOSR-JCE), ISSN: 2278-8727Volume 9, Issue 6.

David,K., Hakob,S. (2011). "Feature Extraction Techniques and Minutiae Based Fingerprint Recognition Process", International Journal of Multimedia Technology, Vol. 1, Issue 1, Pages 31-35.

Chakradhar,K. (2012) Digital Audio Watermarking for Copyright Protection. (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 3 (3) ,4185-4188.