Journal of Engineering and Sustainable Development
www.jeasd.org
Vol. 22, No.01, January 2018
ISSN 2520-0917

EFFICIENT GAIT VERIFICATION THROUGH TEMPLATE PRESERVATION

*Mays Kareem Jabbar Alsabah

Assistant Lecturer, Civil Engineering Department, Misan University, Misan, Iraq.

Abstract: In this paper, the effectiveness of biometric verification system has been achieved with template protection technique. Gait images of person are used; features are extracted from average gait images using Principle Component Analysis (PCA) for dimensionality reduction. At enrollment phase, three gait energy images (GEI)-upper images of person are used. The most reliable component for each person is needed to find and encoded it through Bose Chaudhuri Hocquenghem (BCH) encoding scheme to generate code word. A secrete binary string and hash value are generated for each person and stored in the database along with encoded template and position of strong reliable component. At verification phase, GEI-lower, slow and fast gait images are used. Those features are only selected from original features based on the position of strong reliable component and are used to generate code word and which is decoded using BCH decoder. The hash value is generated and compared with the hash value which was generated in the enrollment phase. If these two hash values are the same then the person is verified otherwise the person is not verified. Also comparison with other recently proposed studies is done to show the effectiveness of the proposed scheme.

Keywords: *PCA, hash value, GEI.*

كفاءة توثيق طريقة المشي من خلال المحافظه على النموذج

الخلاصة: في هذه البحث, تم انجاز كفاءة التحقق من هوية الشخص بتقنية حماية النموذج. لقد تم استخدام صور طريقة مشي الناس واستخراج الملامح من الصور باستخدام تحليل المكون الرئيسي(PCA) لتقليص حجم الأبعاد. في مرحلة التسجيل استخدمت ثلاث صور لطريقة المشي العليا لكل شخص. نحن بحاجه لايجاد عنصر اكثر موثوقية لكل شخص وتشفيرها خلال نظام تشفير ال BCH encoder لتوليد كلمة السر. تم انشاء سلسله ثنائيه وقيمة تجزئه لكل شخص وخزنها في قاعده البيانات مع قالب مشفر وموضع مكون موثوق وقوي. في مرحلة التحقق استخدمنا صور المشي المنخفض البطيئ وصور المشي المنخض السريع. هذه الخصائص فقط التي يتم اختيارها من الخصائص الاصليه اعتمادا على موقف المكون الموثوق القوي. هذه القطع تستخدم لتوليد شفره التي يتم فكها بواسطة BCH decoder. قيمة التجزئه يتم انشائها وتقارن مع قيمة التجزئه المتولده في مرحلة التسجيل. اذا كانت هذه القيمتان متشابهتان يتم التحقق من الشخص وعلى خلاف ذلك لا يتم التحقق من الشخص. بالاظافه الى انه تم القيام بمقارنة نتائج البحث مع نتائج طرق مقترحه مؤخرا لاظهار فعالية الطريقة المقترحة.

1. Introduction

Biometrics means life measurement but the term is usually related with the use of unique physiological characteristics to identify an individual. The most important application of biometric system is related to the security purposes. Psychological report says that people have small but interesting ability of verify the people through their

* m_mays85@yahoo.com

gaits that are known to them [1]. Knowing the person with whom you are communicating is an important part of human collaboration and one expects computers of the future to have the same capabilities. A number of biometric traits have been developed and are used to authenticate the person's identity. The idea is to use the special characteristics of a person to identify him by using special characteristics such as face, iris, fingerprint, and gait [2]. The method of identification based on biometric characteristics is preferred over traditional passwords and PIN based methods for various reasons such as: The person to be identified is required to be physically present at the time-of-identification. Biometric technologies are thus defined as the "automated methods of identifying or authenticating the identity of a living person based on a physiological or behavioral characteristic". A biometric system can be either an "identification" system or a "verification" (authentication) system. Biometric identification system can be used to determine a person's identity even without his knowledge or consent, and biometric verification system can be used to verify a person's identity. Over a last few decades researchers have shown interest and equipped many biometric identifications and verification systems. But these systems have some issues i.e. if biometric information of a specific person which obtained during enrolment phase not protected properly can create the serious security risk for that person. If it reveals once it reveals forever and can't be rescued.

Haruyuki Iwama and Daigo Muramatsu in 2012, presented a gait-based-person verification system which is usually very simple to use that even non expert of gait verification system can use it very easily [3]. Hong Lu, Jonathan Huang, and Tanwistha Saha (2013), produces a system which uses accelerometer sensor and gait characteristic of a person to identify the owners of phone irrespective of any placement of camera and /or device orientation [4]. Holstlaan, and AA Eindhoven in 2004 applied such type of template preservation technique to ear identification and achieved very good results (EER=3%, length of secrete was 100) [5]. Since a gait, specially a biometric verification or identification system provide very useful and attractive help in many security and surveillances situation but it can't be denied the truth that it is a unique, sensitive and unchangeable information about persons.

Because of these issues many people are not enthusiastic to accept system based on biometric authentication. Therefor the stored data should be protected and handled carefully against exploitation [6-9]. This information used to be protected in a proper manner. This study worked on verification of person with a gait verification system even at a distance through cloud network by estimating and protecting his gait patterns.

The rest of paper is arranged as follow: Section 2 presents literature review, Section 3 contains the database, Section 4 contains the scheme of template protection which consists of two phases, Enrollment phase and verification phase. Section 5 presents the proposed scheme results and comparison with other schemes, and the paper is concluded in Section 6.

2. Literature Review

2.1. Biometric verification system

Biometric system is the system which uses the person Biometric systems depends on precise data about matchless biological characters of a specific person in order to work excellently [10]. Depending on the application circumstance, a biometric system typically operates in one of two modes: identification or verification [6].

There are number of feature extraction methods used in literature to extract the suitable feature for different situation. Gabor feature extraction for face recognition extracts only texture information from image [11]. Local binary feature extraction which used for extracting texture information from image and used in age classification and many more field [12-14].

2.2. Principal Component Analysis (PCA)

Principal Components Analysis (PCA) is statistical technique which is used to transform data into new set of coordinates [15, 16]. PCA technique allows finding the data patterns, in such a way that their similarities and differences are determined [15]. Once similarities and differences in data are found, they can compress, to leave the least important dimension. Eigenvalues and eigenvectors are first found from the given data. Eigenvectors represent the direction where the most of our data exists and from the data for PCA in descending order based on eigenvalues.

The eigenvalues represent the variance of data in that direction. Eigenvector with the highest eigenvalue is known as principal component of data. PCA has found its application in many fields such as, image compression, finding data patterns, face recognition system and many more.

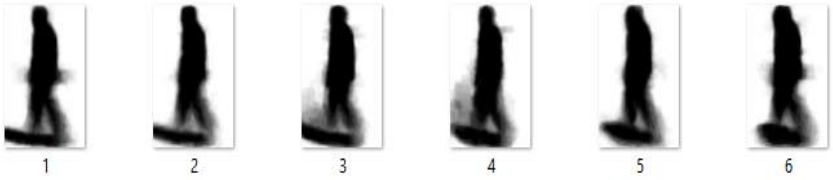
3. Database

Data of 100 persons were used (on concrete and without briefcase) from USF HumanID gait database [17]. Silhouette images of gait sequences of all persons are presented here in a different subfolder. Camera resolution is not given in the documentation of this database and neither is it important because binary images of silhouette frames which have already been preprocessed for noise removal were provided. For each person, a different number of silhouette frames (more than 100) have been provided.

Figure (1): Sample Silhouettes Images of a Person's Gait Sequence

Also inside the gait folder is a text file containing the value of gait period of that person. Using this information, average gait images of each person is formed. As an example, consider that for a certain person, 100 frames and are given the gait period is 20. Then the first 20 gait silhouette images will be averaged together to form the first average gait image, next 20 images would be averaged to form the second average gait image and so on.

(a): Average Gait Images of Person 1.



(b): Average Gait Images of Person 2.

Figure (2): Average Gait Images.

PCA is applied and then quantized feature vector is protected by using cryptographic hash function and stored in data base. First three average gait images of each person are kept for enrollment and the remaining are for verification. At the verification phase strong reliable components of the test image are selected and after decoding, image is compared with data in database. If it is matched with the database sample(s) of claimed identity, then person is verified otherwise the person is not verified. Detailed description of each step followed during this study is mentioned in the following sections.

4. Scheme of Template Protection

This scheme has two phases, Enrollment phase and verification phase.

For each enrolled person, M times of his average gait image for Enrollment are taken.

Basic steps of Enrollment [18] are:

1. Feature Extraction.
2. Statistical Analysis.
3. Quantization.
4. Selecting Reliable Components.
5. Creating Helper Data.

In the verification phase the average gait biometric of a user is captured. Next, feature

extraction and quantization are done. Noise is removed using helper data and secret is reconstructed.

The scheme can be summarized using Figure 3:

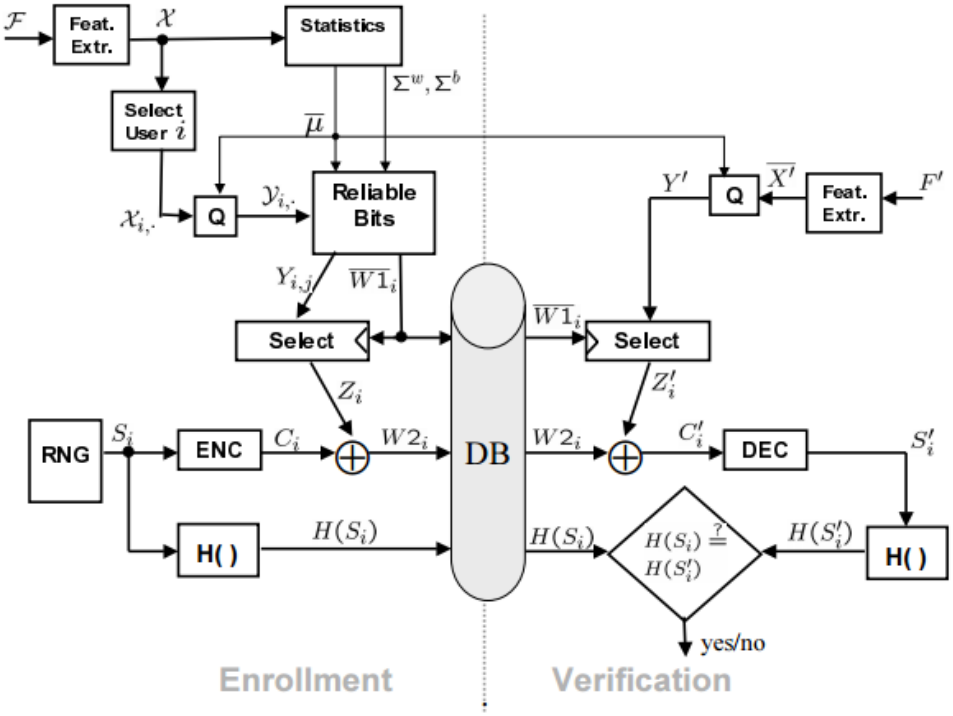


Figure (2) : Basic Scheme [18].

These steps are described in the next sections.

4.1. Enrollment phase

An overview of this stage is presented in Figure 4:

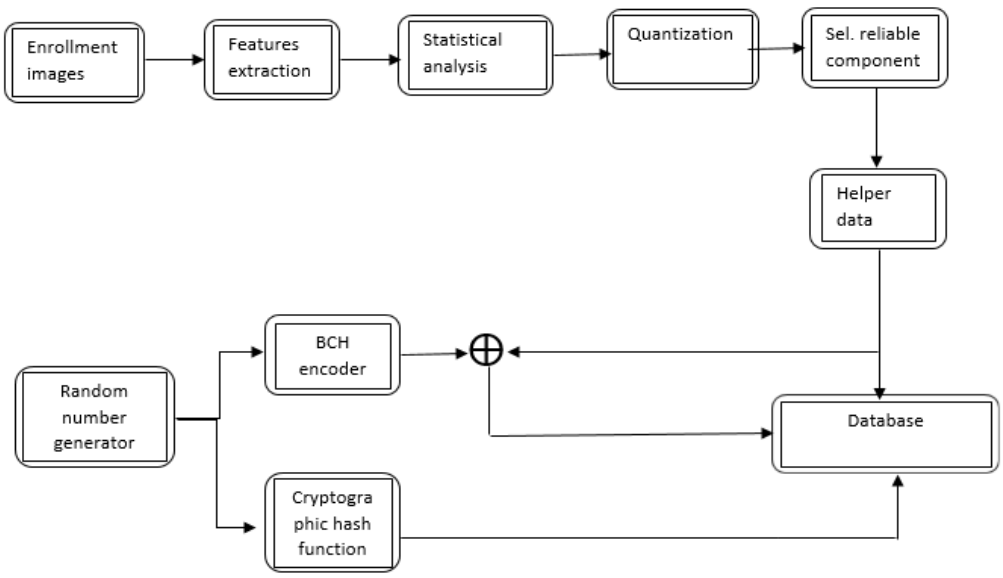


Figure (3): Enrollment phase.

Below is an explanation of all steps represented in Figure 4.

4.1.1. Feature Extraction

PCA feature extraction [19] is quite simple:

i. The mean of all training images in the database are first computed using equation 1.

$$u[j] = \frac{1}{n}\sum_{i=1}^{n} x[i,j] \quad (1)$$

Then to find the feature vector of an image, as it can be seen how it is different from the mean of training data (images) - or how much deviated/shifted are the values of the test images as compared to the training mean. This is simply done by subtracting the mean training image from the test image whose features are required to be computed.

$$shifted\ image = [test\ image - mean\ of\ training\ images] \quad (2)$$

ii. Next, the pixels of the test image as a vector are arranged and find its evector.

iii. Then, the evector is multiplied by the shifted test image vector.

$$feature\ vector\ of\ image = evectors' * shifted\ image \quad (3)$$

This gives us the PCA features.

At the enrollment phase, PCA is applied on the processed average gait images, and reduced dimensionality features of all selected images which have been extracted.

$$X = \{X_{t,r}\}_{t=1\cdots w, r=1\cdots k} \quad (4)$$

Where 't' represents the feature vector, 'r' represents the person. w is number of gait samples per person and k is the number of total persons. X contains w*k features of all images.

4.1.2. Statistical Analysis

In this step, all feature vectors of given persons are calculated by using equation 5 and then the mean values of all mean vectors for all persons are computed by using equation 6.

$$q_r = \frac{1}{k}\sum_{r=1}^{k} X_{t,r} \quad (5)$$

$$m = \frac{1}{w}\sum_{t=1}^{w} q_r \quad (6)$$

Class covariance matrix is calculated by using equation 7, the interior covariance matrix is calculated by using equation 8.

$$Eb = \frac{1}{wk}\sum_{r=1}^{k}\sum_{t=1}^{w}\left(X_{t,r} - q_r\right)\left(\left(X_{t,r} - q_r\right)\right)^T \quad (7)$$

$$Ew = \frac{1}{w}\sum_{r=1}^{w}(q_r - m)\left((q_r - m)^T\right) \quad (8)$$

The mean value of mean feature vector of all feature vectors of data matrix is calculated and the estimate between class covariance and estimates with in class covariance matrix is also calculated.

4.1.3. Quantization

The binary string of the features vector of a person r is calculated by using the feature set of a person $X_r = \{X_{t,r}\}_{t=1\ldots k}$. Then this feature vector is quantized by using the mean value of all feature vectors and then statistical analysis is calculated using equation 9.

$$\left(Q\left(X_{t,r}\right)\right)_t = \begin{cases} 0 & if\ X_{t,r} \le (q)_t \\ 1 & if\ X_{t.r} > (q)_t \end{cases} \quad (9)$$

The quantization process produced binary vectors of all features vectors.

4.1.4. Selection of Reliable Components

As shown in 'Reliable Bits' block in figure 3, the most reliable component for a user in w quantized bits$\left(Q\left(X_{t,r}\right)\right)_{t=1\ldots w}$ is found. For user "r", the "u" most reliable bits that should match to code word length are defined that it will generated in 'creating helper data' block in figure 4.

The most reliable bit can be defined as follows:

For a user "r", a bit is known as a reliable bit if it remains same in all enrollment samples of a person.

For a user if don't get the "u" most reliable bits, the p soft reliable bits is defined of a person as explained in [18]

4.1.5. Formulating helper data

To formulate the helper data, data will be stored in the database and used for the verification of person at verification phase. Helper data comprises of two parts.

1. **First part**, denoted by W1 which stores the positions of the most reliable bits of a person. For a person, the most reliable bits are defined and store their indexes in a vector $W1_r$. And these vectors of all users are stored in the database for future use.
2. **Second part** of the helper data can be found by the following steps:

a. The bits indicated by the vector W1 are placed vector name 'N' for each person 'r'.
b. A bit string of code word length is produced randomly for each person.

c. Results are encoded by using an Error Correcting Code i.e. BCH codes with the specifications (k,s,d) where, 'k' is the length of code word that will generate after encoding, 's' is the length of word that will be encoded, and 'd' gives the information about errors that can be corrected. It will generate the code word, C_r for every person
d. Code word that is generated will be combined with the strong reliable bits, and create second section of the helper data by using equation 10, then store the results in the database.

$$W2_r = N \oplus C_r \quad (10)$$

e. One of cryptographic hash function is used to produce a hash value on the bits generated above and stored in the database.

These two parts of helper data were obtained, i.e. for each person the reliable component was calculated and stored their position in the database. The reliable bits were encoded with the code word generated and stored in the database. Finally, secret keys for all persons were obtained.

4.2. Verification Phase

Basic scheme of verification can be summarized in Figure 5.

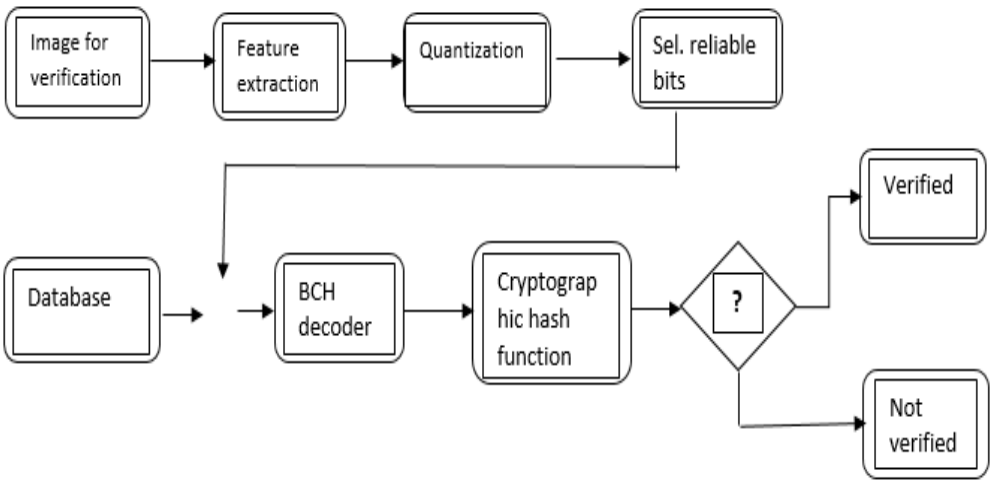


Figure (4): Verification Phase.

For verification:

1. The test image (average gait image) of the person is taken and does preprocessing on it as described above.
2. PCA feature vector X'_r is found and do quantized to obtain binary string $Q(X'_r)$ by equating these values to mean value calculated in statistical analysis block. First section of the helper data '$W1_r$' from the data base is selected and all the bits specified by these indexes are selected from the binary string just derived and are put in the vector 'E'. These most reliable bits are XORed with the

second section of the helper data which stored in the database by using equation 11.

$$C_r = W2_r \oplus E \quad (11)$$

BCH decoder is used to decode the generated code word 'C_r' using equation 11. After decoding, a bit string is obtained on which a one cryptographic hash function is applied to produce a hash value.

This hash value is compared with the hash value which was generated in the enrollment phase. If these two hash values are same then the person is verified otherwise the person is not verified.

5. The Proposed Scheme Results and Comparison with other Schemes

5.1. The Proposed Scheme Results.

In template matching algorithm, a total of 14 strong reliable bits in this experiment are used. Simulation is carried out on MATLAB 2016a.

100 persons from USF HumanID gait database are used [17]. From these 100 persons, each possible binary combination of same and different persons are checked. For example, consider person 1, he will be matched against all 100 persons in the database. Now he should only be accepted when verified against his own identity. When a verification attempt is made against some other person's identity then it should be rejected because that is an imposter attempt.

Each test average gait sample was matched against each enrolled person in the database. Total number of client (genuine user) attempts is calculated and total number of imposter attempts (where actual identity is not the same as the claimed identity) is also calculated.

Then False Acceptance Rate (FAR) and False Rejection Rate (FRR) are computed [20]. Then, Equal Error Rate (EER) is obtained using equation 14. High reliability for the verification system can be obtained when the EER is small. Table 1 shows the percentage values of FAR, FRR, and EER for the proposed scheme.

$$FAR = 100 * number\ of\ false\ acceptances\ /\ total\ imposter\ attempts \quad (12)$$

$$FRR = 100 * number\ of\ false\ rejections/total\ client\ attempts \quad (13)$$

$$EER = (FAR + FRR)\ /\ 2 \quad (14)$$

$$\text{Identification Rate} = 100 - \text{EER} \quad (15)$$

Table 1. Percentage Values of FAR, FRR, and EER for the Proposed Scheme.

Evaluation parameter	Percentage Value
False acceptance rate (FAR)	3.0505 %
False rejection rate (FRR)	4 %
Equal Error Rate (EER)	3.5253 %
Identification Rate	96.4747%

Only 3.5253 % EER and 96.4747% identification rate are obtained. This indicates that the high reliability of the verification system has been achieved.

5.2. Comparison with other Schemes

The proposed scheme achieved 96.4747% identification rate using 100 person images from USF database. The scheme used PCA in the proposed template protection method. The effectiveness is 27% higher than the scheme in [21] that used the same data base. The scheme in [21] adopted the subspace Component and Discriminant Analysis (CDA). The CDA is based on PCA and Multiple Discriminant Analysis (MDA) which seeks to project the original features to a subspace of lower dimensionality.

Meanwhile, the achievement is about 3% higher than the scheme in [22] that used 25 person images only from the USF data base. The scheme in [22] used PCA with Radon Transform (RT). The RT is used to detect features within an image.

Table 2 shows the comparison between the proposed scheme and developed schemes in [21] and [22].

Table 2. Comparison the Identification Rate of the Proposed Scheme with other Schemes.

Schemes	Identification Rate
The Proposed Scheme	96.4747 %
Scheme of [21] 2009	70.63 %
Scheme of [22] 2011	94.23%

5. **Conclusions**

The main idea from this study is to verify a person through his average gait images by protecting its template. This is to make sure that it is stored in the database in an encrypted or protected form, to avoid threat by potential hackers. Biometric systems are quite popular in today's age but their security threats are also not to be neglected. This technique can be applied to any kind of biometric. Experiments on a public gait database were conducted. Only 3.5253 % EER was obtained, which showed that the proposed template protection method with PCA for gait verification is strength.

The resulted identification rate was compared to the other interesting proposed schemes in Table 2 that shows the proposed scheme has high precision.

6. References

1. F. E. Pollick. (2015). *"Psychology of Gait and Action Recognition"*. Encyclopedia of Biometrics, pp. 1280-1285.
2. S. M. J. F. Javier Galbally. (2014). *"Image Quality Assessment for Fake Biometric Detection: Application to Iris, Fingerprint and Face Recognition"*. IEEE Transactions on Image Processing, pp. 710-724.
3. D. Haruyuki Lwama. (2012). *"gait- based person-verification system for*

forensics". Biometrics: Theory, Applications and Systems (BTAS), 2012 IEEE Fifth International Conference, pp. 113-120.

4. J. H. T. S. Hong Lu. (2013). *"Unobtrusive Gait Verification for Mobile Phones"* .ISWC '14 Proceedings of the 2014 ACM International Symposium on Wearable Computers, pp. 91-98.
5. E. V. T. I. D. S. a. T. A. P. Tuyls. (2004). *"Privacy Protected Biometric Templates: Ear Identification"*. Proceedings of SPIE, vol. 5404, pp. 176-182.
6. Y. F. a. B. M. G.I. Davida. (1998). *"On enabling secure applications through off-line biometric"*. IEEE Symposium on Privacy and Security.
7. Y. F. a. B. M. G.I. Davida. (1999). *"On the relation of error-correction and cryptography to an off-line biometric based identification scheme"*. Proceedings WCC99, Workshop on Coding .
8. A. J. a. M. Wattenberg. (1999). *"A Fuzzy Commitment Scheme"* . G. Tsudiked, Sixth ACM Conference, pp. 28-36.
9. J.-P. L. a. P. Tuyls. (2003). *" New shielding functions to enhance privacy and prevent misuse of biometric templates"*. 4th International Conference on Audio- and Video-Based Biometric Person Authentication.
10. *"techopedia"*[Online].Available: https://www.techopedia.com/definition/26990/biometric-system.[Accessed febrarauy 2016].
11. C. L. a. H. Wechsler. (2002). *"Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition"*. IEEE Transactions on Image Processing , vol. 11, no. 4, pp. 467-476.
12. A. H. X. H. M. P. Juha Ylioinas. (2013). *"Age Estimation Using Local Binary Pattern Kernel Density Estimate"*. Image Analysis and Processing – ICIAP, vol. 8156, pp. 141-150.
13. Guoying Zhao. (2011). *"Rotation-Invariant Image and Video Description With Local Binary Pattern Features*". IEEE Transactions on Image Processing , vol. 21, no. 4, pp. 1465-1477.
14. A. H. ,. A. B. E. L. a. R. J. Lotfi Houam. (2014). *"One dimensional local binary pattern for bone texture characterization"*. Pattern Analysis and Applications, vol. 17, no. 1, pp. 179-193.
15. I. Jolliffe. (2002)." *principal component analysis"*. wiely: springer.
16. K. E. a. P. G. Svante Wold. (1987). *"Principal component analysis"*. chemometrics and intelligent laboratory , vol. 2, no. 1-3, pp. 37-52.
17. *"The Data Set"* [Online]. Available: http://figment.csee.usf.edu/GaitBaseline/.
18. P. Tuyls and A. H. Akkerman. (2005). *"Practical Biometric Authentication with Template Protection"*. Springer-Verlag Berlin Heidelberg.
19. B. M. Sabbar and A. A. Al-Sumaily, *"Human Gait Cycle Recognition Using MTI Principles and PCA"*. International Journal of Computer Science Engineering and

Technology(IJCSET).

20. [Online]. Available: https://www.bayometric.com/false-acceptance-rate-far-false-recognition-rate-frr/.
21. B. K, X. T and G. S. (2009). *"Gait recognition using gait entropy image"*. Crime Detection and Prevention (ICDP 2009), 3rd International Conference.
22. J. D. C. A. E. G. M. Hayder Ali. (2011). *"Gait Recognition using Gait Energy Image"*. International Journal of Signal Processing, Image Processing and Pattern Recognition.