Metadata of the chapter that will be visualized in SpringerLink

Book Title	Smart Innovations in Engineering and Technology	
Series Title		
Chapter Title	An Overview of Multi-layer Security System Using Biometrics	
Copyright Year	2020	
Copyright HolderName	Springer Nature Switzerland AG	
Corresponding Author	Family Name	**Chaczko**
	Particle	
	Given Name	**Zenon**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	University of Technology Sydney
	Address	15 Broadway, Ultimo, NSW, Australia
	Email	zenon.chaczko@uts.edu.au
Author	Family Name	**Sethi**
	Particle	
	Given Name	**Harssh**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	University of Technology Sydney
	Address	15 Broadway, Ultimo, NSW, Australia
	Email	harssh.sethi@alumni.uts.edu.au
Author	Family Name	**Alsawwaf**
	Particle	
	Given Name	**Mohammad**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	University of Technology Sydney
	Address	15 Broadway, Ultimo, NSW, Australia
	Email	mohammad.alsawwaf@uts.edu.au

Abstract

After years of research on biometrics, still the authentication and verification of a human being is at risk. Many challenges have been faced by the system to authenticate a human on the basis of their biometric trait. The design of this paper is such that it focuses on three main biometric techniques to extract the various features of a human body contour for authentication and verification purposes. This paper presents an approach towards voice recognition patterns, face geometry patterns and gait pattern analysis to increase the accuracy and precision of the system.

Keywords ■ Multi-layer security, Biometrics, Authentication and verification, Pattern recognition, Machine Vision

An Overview of Multi-layer Security System Using Biometrics

Zenon Chaczko(✉), Harssh Sethi, and Mohammad Alsawwaf

University of Technology Sydney, 15 Broadway, Ultimo, NSW, Australia
zenon.chaczko@uts.edu.au, harssh.sethi@alumni.uts.edu.au,
mohammad.alsawwaf@uts.edu.au

Abstract. After years of research on biometrics, still the authentication and verification of a human being is at risk. Many challenges have been faced by the system to authenticate a human on the basis of their biometric trait. The design of this paper is such that it focuses on three main biometric techniques to extract the various features of a human body contour for authentication and verification purposes. This paper presents an approach towards voice recognition patterns, face geometry patterns and gait pattern analysis to increase the accuracy and precision of the system.

AQ1

Keywords: ■■■ Multi-layer security, Biometrics, Authentication and verification, Pattern recognition, Machine Vision

1 Introduction

Multi-layer security system using biometrics is a system in which multiple biometric characteristics are used to tackle the problem of security systems. The term biometrics has been derived from medieval Greek words, where bio means life and metric means to measure. According to the biometric institute, biometrics is nothing but an automated recognition of an individual based on their particular behavior and biological traits. In general, biometrics is the study of a human being on the basis of their physiological and behavioral traits. As pointed out in the definition given by the biometric institute and various authors, biometrics is further subdivided in two categories, namely, physiological and behavioral [1].

In this contribution, after several authentication schemes and contrasting the existing technology, a model has been presented, which serves to be a reliable model in terms of security. In the following sub sections, various types of biometric characteristics are explained in brief, explaining about the pros and cons of the particular technologies, followed by a discussion on security on the basis of biometrics and finally, the conclusion.

R. Klempous and J. Nikodem (eds.), *Smart Innovations in Engineering and Technology*, Topics in Intelligent Engineering and Informatics 15,
https://doi.org/10.1007/978-3-030-32861-0_15

2 Biometric Characteristics

Face: Face geometry is physical feature of a human body [1]. Face has many characteristics to consider while comparing such as the width of distance between the eyes, width of the nose, length of jawline etc. This trait can be used for the purposes, authentication and verification. For this trait, some selected facial features are selected for verification and authentication purpose. Initially, the features are extracted from the image and compared with other images.
Voice Pattern: Voice recognition patterns have been introduced in the recent years and have proved to be a reliable source for security [4]. This biometric technique is a behavioral technique as it considers many factors and no new hardware is required because this technology only needs a microphone to record the words or phrases. Speaker recognition is another technique to identify a human being with the help of their speaking abilities/ characteristics.
Gait: Gait analysis is a physiological attribute between the two biometric types [5]. This type of technique focuses on the walking style of a human being. The main advantage of this technique is that it is non-invasive compared to other biometric characteristics. There are various factors to consider while performing a gait analysis, such as the stride length, stride time, the bend in knee etc. this technique involves the walking of a human being, which is converted in to a 2D image and then utilized for further analysis.

3 Methdology

It has been observed that all the biometric techniques are different as they all depend on various factors. For each technique, there are different variables. These variables are used to identify and authenticate the particular person. In this contribution, three main techniques have been utilized.

Initially, for any individual to give access, in terms of security, that particular individual has to go through a series of events by authenticating the biometric characteristics. In this paper, voice patterns have been chosen as the first biometric characteristic followed by the face geometry patterns and finally the Gait analysis.

3.1 Voice Recognition

A voice recognition system has four main domains, which are: (1) Extraction of the feature, (2) acoustic modeling, (3) pronunciation modeling and (4) decoding [4]. Initially, the user has to speak in to the microphone creating sound waves. These sound waves are then converted into electrical signals in a digital form in order to be transmitted. These digital signals are transformed so that the system can understand them. Further, these signals are converted into a form of vectors, which only contains the relevant information regarding the users voice. Finally, all the words spoken by the user are transmitted into a digital form and stored in the database with proper classification. However, at times, the information

extracted from these signals can be less discriminative, therefore, an extraction model has been used to provide avoid the discrepancies. This feature extraction model is the Mel Frequency Cepstral Coefficient (MFCC).

MFCC is one of the various feature extraction techniques used for the purpose of voice recognition. In this technology, initially, different speech samples are collected from different speakers and converted into a digital form by recording these signals at regular intervals. This data of digital form is then quantized in order to eliminate the noise. And finally, these samples of speech are tested for feature extraction purposes. Feature extraction is a function, in which the signal is reconstructed with the presence of original data.

3.1.1 Decoding This step is considered as the most important step as it is performed to search for the best match for the retrieved speech samples from the speech database. In this step, a decoder is used to make the ultimate decision based on combining and optimizing the data retrieved from the acoustic and language models of an individual.

Acoustic modeling has two kinds, namely, word model and phoneme model. HMM (Hidden Markov Model) is the most common form of acoustic modeling, which is used for speech recognition system [4]. There are a few states in HMMs, which vary on the output probabilities with a finite number of states. In HMM, one state is responsible for the modeling of the variability in speech known as the transition probability, whereas, the other processes are more focused on the output formulated by Gaussian mixture. In this model, the previous state is responsible for the probability of being in any state by using the Viterbi algorithm.

3.1.2 Pronunciation Modeling This step is initiated during the recognition process when the sequence generated by HMM is compared with the words of dictionary to formulate the final output containing the information about the words/phrases known to the system. Decoder is then used to recognize the spoken words by combining the information retrieved from acoustic models.

3.1.3 Approaches Made Towards Voice Recognition For voice recognition pattern system, there are three different type of approaches made, namely: Acoustic phonetic, pattern recognition and artificial intelligence approach.

Acoustic phonetic approach, as the name suggests, utilizes the information of linguistics and phonetics to guide. This process has a poor performance as it finds difficult to express the rules and to improve the system. This type of approach identifies the words, structure of sentence, meaning or the individual phonemes.

The next approach is the pattern recognition approach, which has two steps. The first step involves the training of speech patterns, whereas, the next step involves the recognition of these patterns. A test pattern is created on the input signal, which tests the various other patterns with the reference pattern and test the similarity between the test pattern and the reference pattern.

The last approach is the artificial intelligence approach. This type of approach is a skilled system implemented to classify different sounds. It is a combination of both, acoustic approach and pattern recognition approach.

The Voice recognition pattern analysis is the initial technique used for the authentication and verification of an individual, which is followed by other techniques.

3.2 Face Geometry Analysis

Face geometry analysis of a human face depends on various factors such as the width of the nose, distance between the eyes etc. In order to distinguish a human being from another, relative location of face objects such as the face length, width, and height are considered, as they are unique for every individual [5]. The most common advantages of using face geometry are:

- Low cost algorithm.
- Medium cost camera to capture images.
- Low storage images, as the template size is low.
- User-friendly, as the system is easy to access compared to other biometric techniques.

In the Fig. 1, a step by step instructions are given as to how the face geometry analysis works, starting from capturing the image to converting the image to a 2D image to compare the features of the face.

According to the image captured, first the distance between the eyes is calculated, which is depicted in the image A. After the distance between the eyes is calculated, the distance between the ears is measured in image B. Similarly, in the image C, the distance from the forehead to nose is calculated, while in image D, the leap in the human face is calculated. After all the distances are measured, angles are calculated for further accuracy. These angles are made between the eyes and the nose, which can be seen in image E. Similarly, to increase the efficiency, the angles between the ears and the mouth are measured using a triangle, shown in image F. Both the images, E and F, have a triangle, in which the sum of angles is 180.

The main advantage of this technique is that the image captures the human face in any condition, be it dry weather or a rainy weather, the processor focuses on the human face. As there are no harmful effects of using the processor, this technique is considered to be efficient.

In the proposed methodology, the image is captured via a digital camera and transformed into a binary image. Since the image obtained is in binary, therefore, it presents itself in a black and white image, as seen in image F.

The distance to be measured can be done using MATLAB based on counting the pixel distances. Once the center points from the face areas are captured, the distance is then calculated using the cosine distance, Chebychev distance and the Euclidean distance.

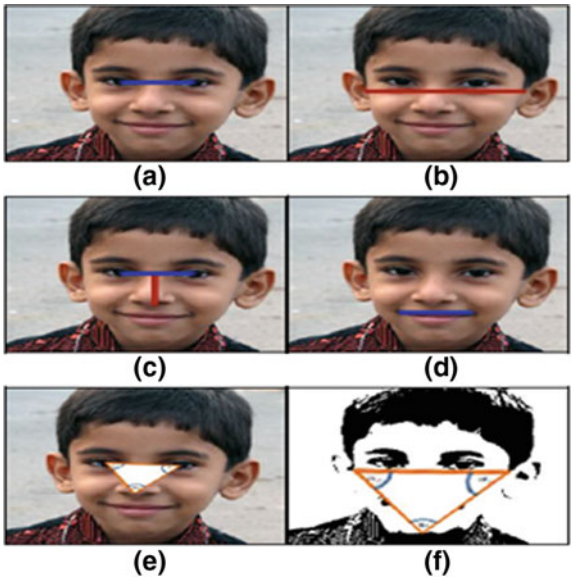

Fig. 1. Face feature analysis.

3.2.1 Mathematical Analysis In order to achieve the maximum efficiency of face geometry analysis, three main metric distance and angle measurement techniques are used. These techniques are:

1. *Cosine distance.*
   Once the image is in 2D, cosine distance is calculated using the distance between two points a and b, where p and q have coordinates and $\cos(p, q)$ is the distance to be calculated according to the equation shown below.

$$\cos(W_1, W_2) = \frac{\sum_{i=1}^{n} P_i Q_i}{\sqrt{\sum_{i=1}^{n} P_i^2 \sum_{i=1}^{n} Q_i^2}} \tag{1}$$

   Salton and Buckley [5] explain the concept of this equation in detail. In accordance with this technique, the distance between nose, mouth and eyes of a human face can be calculated.

2. *Chebychev distance.*
   This technique involves the same parameters as the cosine distance. According to this technique, the Chebyshev distance is $D(p, q)$. The equation is:

$$D(p, q) = max_i(|p_i - q_i|) \tag{2}$$

   The Chebyshev distance is an equation, which is induced by uniform norm or supermum norm [4]. Similar to Cosine function, this equation calculates the distance between the eyes and the nose.

3. *Euclidean distance.*
The Euclidean distance [2] is the ordinary distance between two points. As we recall the Pythagorean formula, the Euclidean formula is based on it. The Euclidean distance equation is shown below.

$$d(p,q) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \cdots + (q_n - p_n)^2} = \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2} \quad (3)$$

The $d(p,q)$ is the length and in the equation shown above, n is the dimension of feature vector, whereas, p_i and q_i are the two different components of face. The Euclidean distance is generally used to calculate the distance between two positions of a human face.
In this contribution, three different techniques have been used to determine the face geometry of a human being based on the various factors. The factors to consider in measuring the face geometry were the eyes, nose and the mouth of an individual. Three different techniques have been used in order to increase the efficiency of the system as each technique focuses itself on different feature of the face. Face geometry is a technique in which very less equipment is required for the authentication and verification of an individual.

3.3 Gait Analysis

Gait can be defined as the periodic motion of a human body, which can be predicted. According to Murray [3], a humans periodic motion is a total walking cycle. It has been observed that the pattern of the walking style of any individual is similar to the sinusoidal nature of a graph. However, the gait analysis is a system based on the relationship between the body segments and the joint angles [6]. In the recent past, many researches have been carried out on various body segment-links to determine the joint points.

In order to extract the body points, position of chest, shoulder, knees, ankles, pelvis and neck are considered. In this technique, there are two parts. First part deals with the upper body region, whereas, the second part focuses on the lower body region. Upper body region includes shoulder, neck, head and chest. For the upper body region, the horizontal coordinate is measured according to the following equation:

$$X_{center} = X_s + (X_1 - X_s)/2 \quad (4)$$

According to the equation, X_s and X_1 corresponds to the first and the last position of pixel on scan line. The coordinates of shoulder and chest are responsible for the waist and pelvis coordinates, which are calculated by the interpolation formula shown below.

$$X_{waist} = X_{pelvis} = X_{chest} + \frac{(x_{chest} - X_{shoulder})}{(Y_{chest} - Y_{shoulder})}(Y_{waist} - Y_{chest}) \quad (5)$$

The upper body region components are much easier to find and the movement of these components is minimal, whereas, the knees and ankles require an algorithm focusing on the leg angles from the body contour by linear regression analysis. In order to extract the data, anatomical data is utilised to measure the position of knees, ankles and pelvis. In the Figs. 2 and 3 shown below, the first shows the position of the joints and the border pixels extracted from the body contour, whereas second shows the extracted data from Fig. 2 with lines as plotting on a graph.

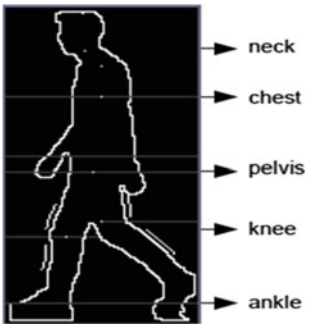

Fig. 2. Gait analysis (A).

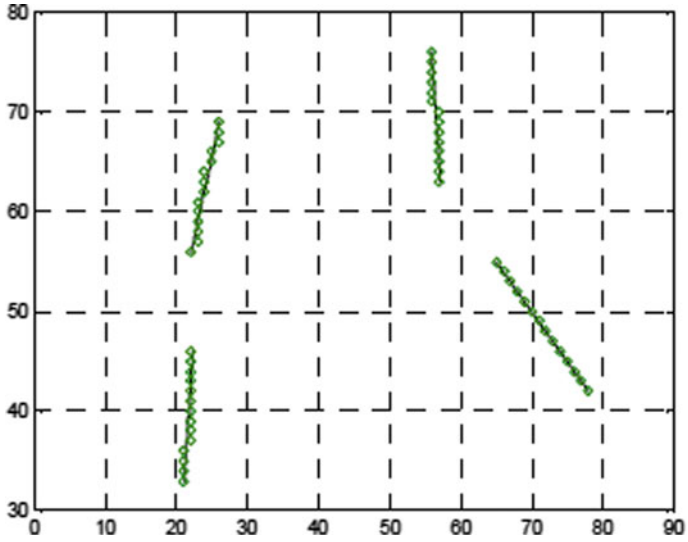

Fig. 3. Gait analysis (B). Measurement unit of height and distance in cm.

The images used in Figs. 2 and 3 can be utilised to mark the joint positions on a human body contour. Linear regression can also be used to calculate the thigh and shin angles. In order to calculate the angle at a joint in a frame k, the following equation is used:

$$\theta_{jnt,k} = \sum_{i=1}^{n}(y_i\overline{y})(x_i - \overline{x}) \Big/ \sum_{i=1}^{n}(x_i - \overline{x})^2 \tag{6}$$

According to the above equation, n is assumed to be the number of points and $jnt \in (shin, thigh, knee)$. There is a certainty factor for the angle $\theta_{jnt,k}$ which can be measured by the following formula. In Eq. 7, $Y_{jnt,k}$ is the certainty factor.

$$Y_{jnt,k} = \left| \sum_{i=1}^{n}(y_i - \overline{y})(x_i - \overline{x}) \Big/ \sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2(x_i - \overline{x})^2} \right| \tag{7}$$

Any image that is captured can be susceptible to noise, which can lead to spurious pixels and further contaminate the data at hand. Therefore, in order to reduce the noise in the image, a weighted moving average is applied, which can be seen by the following equation. In the equation, q is assumed to be a moving window size, which is set to 2.

$$\theta_{jnt}(n) = \sum_{k=n-q}^{n+q}(y_{int,k} \cdot \theta_{jnt,k}) \Big/ \sum_{k=n-q}^{n+q} y_{int,k} \tag{8}$$

The following equation determines the knee and ankle points, where ϕ is the phase shift, L_s defined as the length of the shin and x_i and y_i are assumed to be the joint coordinate of pelvis.

$$X_{ka}, Y_{ka} = [X_i L_s \cos(\theta_{jnt} + \phi) Y_i L_s \sin(\theta_{jnt} + \phi)] \tag{9}$$

In the analysis of gait, the factors to consider are the linear and angular position of the joints of a human body contour. Using these points and factors, results has been calculated.

Biometrics has been proved to be a reliable source for decades, but in this contribution, an attempt by utilizing three different techniques to create a hybrid approach is achieved for the better future. Initially, voice recognition pattern to identify a person, followed by the face recognition for authentication and using the same camera to confirm the individuals identity by using gait analysis is used in this contribution.

4 Summerizing the Results

In order to summarize the results for the proposed hybrid approach of combining the three techniques together, each technique have been performed with different

apparatus and different surroundings. Therefore, for each technique, the results formulated are separate.

To start with the results, the first technique is the voice recognition pattern analysis. In this technique, the users were asked to speak words out of the list for a few times. Subsequently, the users were asked to speak the same words randomly to recognize the patterns.

In the following Table 1, the results have been showed, in which there are 2 male and 2 female speakers. All the speakers have been experimented in front of a computer, human being, robot and a mobile.

Table 1. Experiment Results of 4 speakers, 2 male and 2 female. Inaccuracy range $-+0.5\%$.

	Speaker1	Speaker 2	Speaker3	Speaker4
Robot	90	80	90	100
Human Being	100	90	100	100
Computer	100	100	100	90
Mobile	100	100	90	100
Average (%)	**97.5**	**92.5**	**95**	**97.5**

According to the results formulated, it can be observed that the overall efficiency of the speaker identification system is approximately 96%. However, In Table 2 shown below, it is observed that the speech recognition system has an efficiency of approximately 98%.

Table 2. Overall efficiency of speaker identification Inaccuracy range $-+0.5\%$.

	Recognition (%)
Robot	98
Human Being	100
Computer	97
Mobile	97
Average (%)	**98**

After gathering the results from the voice recognition patterns, we assembled the results from face recognition patterns.

In order to perform the analysis, several images of face were taken. Features were extracted in the procedure explained in Sect. 3.2 "Face Geometry Analysis" of this chapter. Once the features are extracted from the face recognition patterns, the image is stored in the database as a template, which could be accessed by the administrator. In this approach, along with the extraction of features of the face, many other things have been considered such as the angle representation technique to increase the efficiency of the system.

From past studies, the biometric systems accuracy depends on two metrics, namely, False Rejection Rate (FRR) and False Acceptance Rate (FAR). FRR is basically the rejection of individuals by the system even though they are the correct ones, on the other hand, FAR can be defined as the percentage of unauthorized people getting the access due to an error in the system. Another case, when both, FRR and FAR, have equal values, it is often referred to as Equal Error Rate (ERR). In the following Table 3, the techniques used for face recognition patterns have been observed and the results have been noted for their success and failure of that particular technique.

Table 3. FRR and ERR for face recognition patterns. Inaccuracy range $-+0.1\%$.

Process Used	FRR	ERR
Cosine distance	0.0	0.0
Euclidean distance	0.0	0.0
Chebyshev distance	0.0	0.0
Parameter of angles	0.0	0.0

In experimenting, a computer system along with a digital camera was used. The best processes to calculate the face recognition pattern were used with the best FRR, FAR and ERR values, resulting in 100% efficiency of the system as both, FAR and FRR values were observed to be 0%. The techniques were tested in the best way possible with appropriate surroundings.

The last biometric technique is the gait pattern analysis. This technique required a lot of analyzing as the technique had many equations to consider. In the results observed for gait pattern analysis, many consecutive images were observed for a sequence in order to analyze the results. From the data acquired, the following figure explains the various results observed while performing the experiment.

In Fig. 4, the signatures of gait have been considered for three people during a gait cycle, focusing on the angles of thigh and knee.

Similarly, to focus further on the knee and the thigh angles in a gait cycle, Figs. 5 and 6 explain the knee and thigh angles according to the time of a gait cycle.

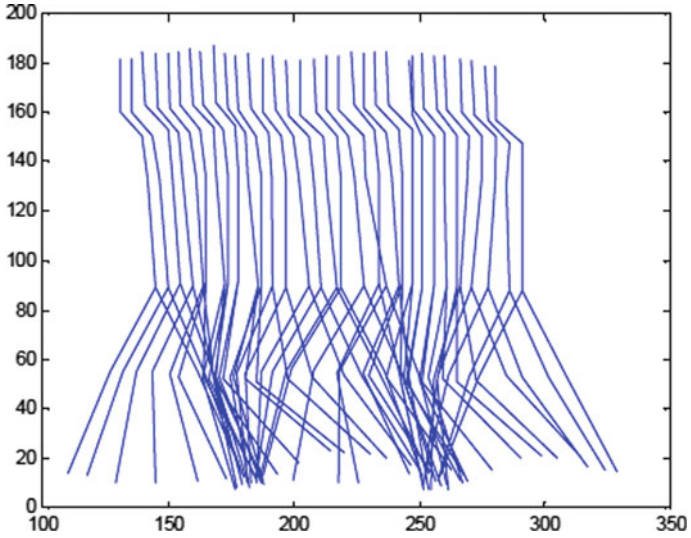

Fig. 4. Signature of gait for three people overlapping in a gait cycle. Measurement unit of hight and distance is cm.

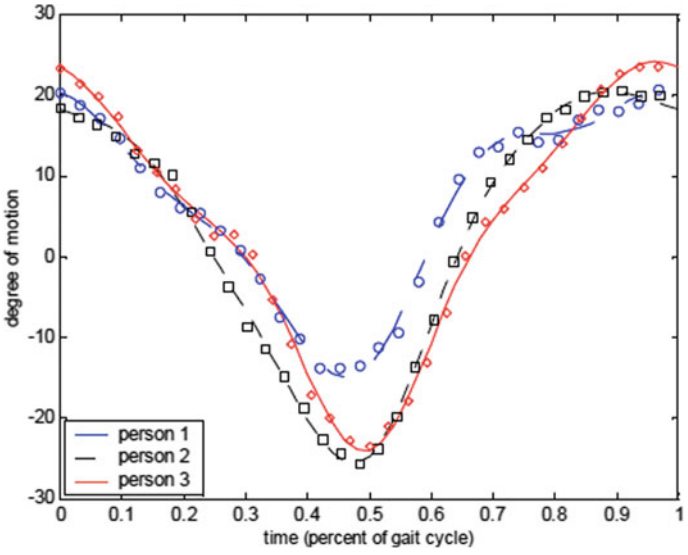

Fig. 5. Thigh angles.

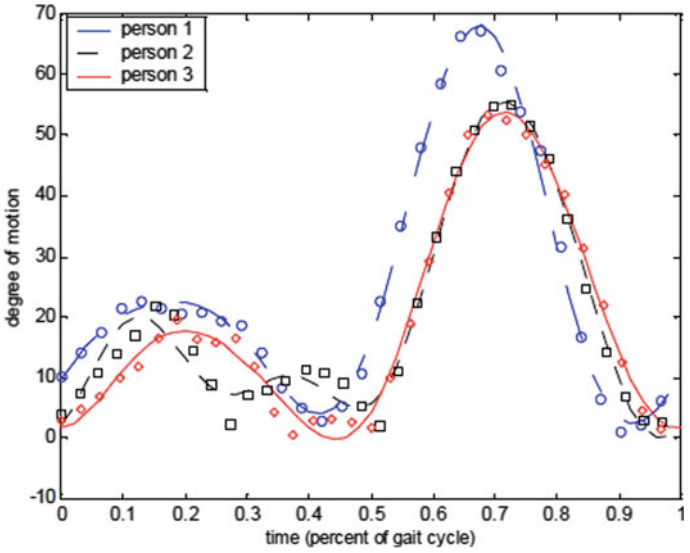

Fig. 6. Knee angles.

In Fig. 7, the angular displacement of a human body contour can be observed, which is periodical in nature. The observed results were predicted using a phase-space portrait.

In order to recognize a human, back-propagation algorithm has been implemented, in which only the kinematic features have been considered. Although the experimental data was limited and the results were a 100% success, but with the large database, the system will perform with the desired results.

The results in this contribution have provided a meaningful comparison for future and help extend the analysis further.

5 Conclusion

We have described a new hybrid approach for extracting the results from three different biometric techniques in order to try and achieve the maximum level of security. It was challenging to deal with three complex techniques. In this paper, an attempt has been made towards how much the technology has innovated in the recent years. Voice recognition pattern is an interesting area to study as humans recognize people on the basis of their biometric traits. However, in distinguishing between various users for a machine, this approach has been successful. We have explained the authentication process based on utilization of knowledge from hand geometry, KLT (Karhunen Loeve Transform) coefficients to determine the face detection, and recognition patterns. A significant amount of research studied these techniques, which has helped to introduce a hybrid approach of utilizing the three biometric approaches. It can be implemented from the results that the

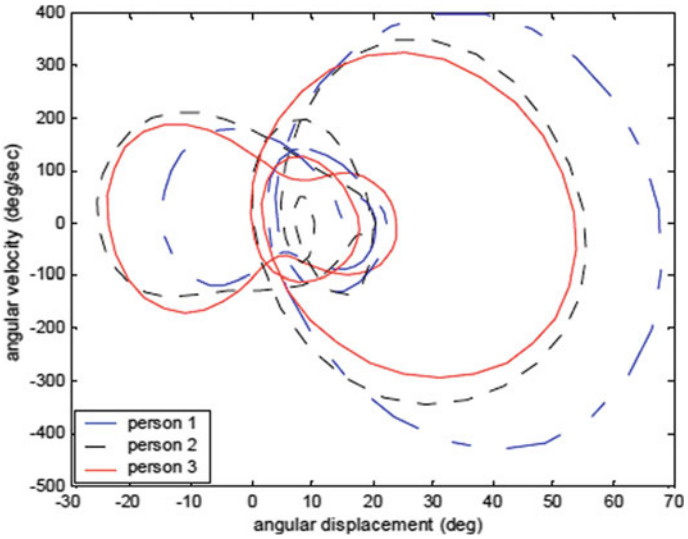

Fig. 7. Phase-shift portrait.

techniques used to authenticate and verify an individual have been successful to reach an efficiency of approximately 99%, using all the three techniques. In this contribution, one of the main purposes of this was to implement an approach, in which the weakness of one technique is considered as the strength of the other. In addition, to further enhance the performance of the system, various other biometric technologies can be used such as the nose and ear geometry, and minutiae extraction from fingerprint etc.

References

1. Brunelli, R.: Template Matching Techniques in Computer Vision: Theory and Practice. Wiley (2009)
2. Deza, M.M., Deza E.: Encyclopedia of distances. In: Encyclopedia of Distances, pp. 1–583. Springer, Berlin (2009)
3. Murray, M.P., Drought, A.B., Kory, R.C.: Walking patterns of normal men. JBJS **46**(2), 335–360 (1964)
4. Saini P., Kaur P.: Automatic speech recognition: a review. Int. J. Eng. Trends Technol. 132–136 (2013)
5. Swamy, S., Ramakrishnan, K.: An efficient speech recognition system. Comput. Sci. Eng. **3**(4), 21 (2013)
6. Yoo J.H., Nixon M.S., Harris, C.J.: Extracting gait signatures based on anatomical knowledge. In: Proceedings of BMVA Symposium on Advancing Biometric Technologies, London, UK, pp. 596–606 (2002)

Author Queries

Chapter 15

Query Refs.	Details Required	Author's response
AQ1	As keywords are mandatory for this chapter, please provide 3–6 keywords.	

Multi-layer security,
Biometrics,
Authentication and verification,
Pattern recognition
Machine Vision

MARKED PROOF

Please correct and return this set

Please use the proof correction marks shown below for all alterations and corrections. If you wish to return your proof by fax you should ensure that all amendments are written clearly in dark ink and are made well within the page margins.

Instruction to printer	*Textual mark*	*Marginal mark*
Leave unchanged	··· under matter to remain	New matter followed by
Insert in text the matter indicated in the margin		or (x2)
Delete	through single character, rule or underline or through all characters to be deleted	or (x2)
Substitute character or substitute part of one or more word(s)	through letter or through characters	new character or new characters
Change to italics	under matter to be changed	
Change to capitals	under matter to be changed	
Change to small capitals	under matter to be changed	
Change to bold type	under matter to be changed	
Change to bold italic	under matter to be changed	
Change to lower case	Encircle matter to be changed	
Change italic to upright type	(As above)	
Change bold to non-bold type	(As above)	
Insert 'superior' character	through character or where required	or under character e.g. or
Insert 'inferior' character	(As above)	over character e.g.
Insert full stop	(As above)	
Insert comma	(As above)	
Insert single quotation marks	(As above)	or and/or or
Insert double quotation marks	(As above)	or and/or or
Insert hyphen	(As above)	
Start new paragraph		
No new paragraph		
Transpose		
Close up	linking characters	
Insert or substitute space between characters or words	through character or where required	
Reduce space between characters or words	between characters or words affected	