Fusion of Face Recognition and Facial Expression Detection for Authentication: A Proposed Model

Delina Beh Mei Yin
Malaysian Institute of Information Technology
Universiti Kuala Lumpur
1016 Jalan Sultan Ismail
Kuala Lumpur
delina@unikl.edu.my

Shariman Omar
BT Global Technology
Menara BT, Tower 3
Avenue 7, Bangsar South
Kuala Lumpur
shariman.omar@bt.com

Bazilah A. Talip
Malaysian Institute of Information Technology
Universiti Kuala Lumpur
1016 Jalan Sultan Ismail
Kuala Lumpur
bazilah@unikl.edu.my

Amalia Muklas
Malaysian Institute of Information Technology
Universiti Kuala Lumpur
1016 Jalan Sultan Ismail
Kuala Lumpur
+603 2175 4355
amalia@unikl.edu.my

Nur Afiqah Mohd Norain
Malaysian Institute of Information Technology
Universiti Kuala Lumpur
1016 Jalan Sultan Ismail
Kuala Lumpur
afiqahmdnorain@gmail.com

Abu Talib Othman
Malaysian Spanish Institute
Universiti Kuala Lumpur
Kulim Hi-Tech Park,
Kedah
abutalib@unikl.edu.my

ABSTRACT

The paper presents a novel model of hybrid biometric-based authentication. Currently, the recognition accuracy of a single biometric verification system is often much reduced due to many factors such as the environment, user mode and physiological defects of an individual. Apparently, the enrolment of static biometric is highly vulnerable to impersonation attack. Due to the fact of single biometric authentication only offers one factor of verification, we proposed to hybrid two biometric attributes that consist of physiological and behavioural trait. In this study, we utilise the static and dynamic features of a human face. In order to extract the important features from a face, the primary steps taken are image pre-processing and face detection. Apparently, to distinguish between a genuine user or an imposter, the first authentication is to verify the user's identity through face recognition. Solely depending on a single modal biometric is possible to lead to false acceptance when two or more similar face features may result in a relatively high match score. However, it is found the False Acceptance Rate is 0.55% whereas the False Rejection Rate is 7%. By reason of the security discrepancies in the mentioned condition, therefore we proposed a fusion method whereby a genuine user will select a facial expression from the seven universal expression (i.e. happy, sad, anger, disgust, surprise, fear and neutral) as enrolled earlier in the database. For the proof of concept, it is proven in our results that even there are two or more users coincidently have the same face features, the selected facial expression will act as a password to be prominently distinguished a genuine or impostor user.

CCS Concepts

• **Security and privacy~Biometrics**

Keywords

Biometric, face recognition, facial expression, authentication, identity verification.

1. INTRODUCTION

Traditional authentication system such as passwords and passcodes are not reliable which can be lost, stolen or disclosed. As mentioned by [1], the most common issues for passwords security are the length and memorability. The length of a password determined on how easy for the perpetrator or program to break the encryption. However, by adding more characters also does make it more difficult to remember which tends to make people write it down. This defeats the purpose of security as another person who finds it out would not need to perform any process of cracking it.

Nowadays, data theft has been rampant within the renowned organization. By and large, people are using passwords and passcodes as a basic data protection. Unfortunately, such conventional method suffers from several dilemmas and challenges such as memorability issue when a password is adequately complex or prone to exploitation due to lack of passwords' complexity. As a result, the emergence of biometric is to replace passwords as it is much easier to use. However, it does not mean that biometric is fully secured as every system has its own benefits and drawbacks. Nevertheless, by using biometric it does not mean that one is completely secured from any attacks. Subsequently, with the technology nowadays by having only one layer of authentication, it is easily bypassed by the perpetrator. With that in mind, adding another layer of authentication is recommended as it will be harder and takes a longer time to bypass it.

IMCOM '17, January 05-07, 2017, Beppu, Japan
© 2017 ACM. ISBN 978-1-4503-4888-1/17/01…$15.00
DOI: http://dx.doi.org/10.1145/3022227.3022247

In the recent years, biometric becomes a well-known alternative in identity verification because every individual's biometric is unique, cannot be stolen and most importantly cannot be forgotten. Earlier literatures [2, 3] documented that a single modal (unimodal) biometric used in identity verification has posed several limitations as follows:

(a) Noise in sensed data: Noisy data could result from unfavourable ambient conditions, (e.g., poor illumination of a user's face in a face recognition system)

(b) Intra-class variations: These variations are typically caused by a user who is incorrectly interacting with the sensor (e.g., incorrect facial pose)

(c) Non-universality: Some people cannot physically provide a standalone biometric credential due to illness or disabilities, thus the biometric system may not be able to acquire meaningful biometric data from a subset of an individual.

(d) spoof attacks: An intruder attempts to masquerade as someone else by falsifying data and thereby gaining illegitimate access and advantage [4].

Therefore, limitations and weaknesses of single modal biometric systems [5, 6] can be resolved using multimodal biometric. Recently there has been a lot of interest in multimodal verification systems in which two or more sources of biometric attributes are utilised for verification purposes. For multimodal-based biometric recognition, a lot of studies and researches have been done at different levels of fusion. As for an instance, the authors in [7, 8] have proposed fingerprint and finger vein identity authentication system based on multi-route detection. As such, their experimental results of the fusion scheme have successfully overcome the limitation of single fingerprint or single finger vein recognition respectively [7, 8]. The aforementioned integrated approaches had made it difficult for an imposter to spoof both traits simultaneously. Peng et al. [9] proposed a score-level fusion approach for finger multimodal biometric authentication that combines finger vein, fingerprint, finger shape and finger knuckle print features of a single human finger. As a result, the fusion approach by Peng and his associates [9] obtains a larger distance between genuine and imposter score distribution as well as achieved lower error rates. Apart from fusing the attributes of the finger for identity verification, there are other researchers proposed the integration of different modal biometric. The work conducted by [10] proposed an identity verification method by combining face and iris biometrics using Fisher's discriminant analysis and a neural network with radial basis function to classify the vector as being genuine or an impostor. Determan et. al [11] has filed a patent for their research entitled combined face and iris recognition system.

In this study, we proposed an integration of face recognition and facial expression detection for verifying identity and authentication. The reason of using human faces in this study is because human faces contain abundant information of human facial behaviour, therefore human face play an important role in social communication. Besides, face biometric is also widely used in surveillance, security authentication, forensic and perhaps other commercial application. As compared with other biometrics systems using fingerprint or palm print and iris as mentioned earlier, face biometric has distinct advantages because of its non-intrusive physiological attribute and non-contact process. Face images can be captured from a distance without touching the person being identified, and the identification does not require interacting with the person. As a result, face biometric demonstrates the interest in identifying humans in surveillance scenarios, security access control etc. [12]. Meanwhile, unique expression of a human pragmatically contributes to intelligent security system for real time surveillance [13] security authentication [14] as well as emotion recognition [15].

Motivated from the aforementioned existing literatures and security applications, we proposed to solely utilise face biometric by combining both physical and behavioural biometric attributes of a human, which strengthens its system to overcome the flaws of each other. The first authentication is where the proposed system is to identify the genuine user. By integrating the facial expression as a second authentication, an additional layer of verification is added to detect whether the facial expression is owned by the genuine user who has passed the first authentication. For an instance, consider there are two people who coincidentally have a similar facial appearance. In this case, there is a potential risk of a false accept for a system based purely on face recognition as a relatively high match score may be achieved when matching one against the other. However, if the same system also included facial expression as a unique key, it would be very unlikely that any given two people would have similar faces and similar selected facial expression. Evidently, the ability of the system to distinguish between people is increased significantly.

2. FACE DETECTION USING HAAR-LIKE FEATURES

Although there are many different existing algorithms to perform face detection, each has its own weaknesses and strengths. As for an instance, some used skin tones and contours, and some used even more complex involving templates, neural networks, or filters. These algorithms suffer from the same problem; they are computationally expensive [16]. A face image acquired from the user is only a collection of colour and/or light intensity values. Due to the wide variations of shape and pigmentation within a human face, analysing these pixels for face detection is time consuming and difficult to accomplish. Pixels often require reanalysis for scaling and precision. Therefore, Viola and Jones devised an algorithm, called Haar Classifiers, to rapidly detect any object, including human faces at the pre-processing stage, using AdaBoost classifier cascades that are based on Haar-like features and not pixels [17, 18]. A Haar-like feature considers neighbouring rectangular regions at a specific location in a detection window, sums up the pixel intensities in each region and calculates the difference between these sums. This difference is then used to categorise subsections of an image. For an example, the areas around the eyes are darker than the areas on the cheeks. Later, Lienhart and Maydt made an improvement by adding rotated features at 45 degree, significantly improving the classifier performance [19] when applied to objects with diagonal shapes. In this case, the improved Haar-like features particularly well-suited to our proposed work. There are a total of 14 feature prototypes developed by [19] as shown in Figure 1 which include 4 edge features, 8 line features and 2 centre-surround features. Haar-like features, as shown in Figure 1 are used to detect a face image. Haar-like features can be independently scaled in vertical and horizontal direction to allow these features to be used to detect objects of various sizes.

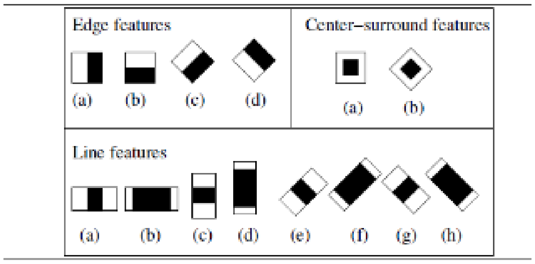


Figure 1. Examples of the used Haar-like features in face detection process [19].

A cascade of classifiers is a degenerated decision tree where at each stage a classifier is trained to detect almost all objects of interest (frontal faces in our example as shown in Figure 2) while rejecting a certain fraction of the non-object patterns [18]. The work in [18] constructed a cascade of classifiers which is able to increase detection performance while radically reducing computation time. The system detects objects in question by moving a window over the image. Each stage of the classifier labels the specific region defined by the current location of the window as either positive or negative. In this case, positive means that a human face is identified whereas negative means that a human face cannot be found in the image.

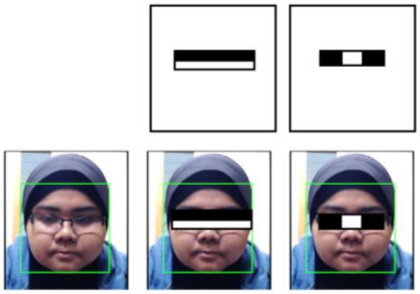


Figure 2. Face Detection using Haar-cascade algorithm

As shown in Figure 2, a user will pose his or her frontal face in front of a camera. During the process, a square frame will appear and detect the face area of the person. It shows the process of face detection. The cascade is performed as follows: first feature (as depicted in Figure 2) is tested and if the result is positive, the algorithm goes to the next one, otherwise it ends and no face is detected. The process is repeated for all features in Figure 1. They are sorted from the most distinctive to less distinctive; therefore, classification of non-faces can be rapidly done because the tested region is excluded in first steps of the cascade.

3. FACE RECOGNITION

In the facial recognition process, a person from a digital image or a video frame from a video source can be identified and verified by comparing selected facial features from the image and a facial database. Therefore, after the face has been detected and located in the image or video frame during the pre-processing, it can be analysed in terms of facial action occurrence. In our work, during the enrolment process, all users' details will be stored in the database as required. The user is prompted to capture a set of face images with different poses included seven universal face expressions (i.e. neutral, happy, sad, angry, fear, surprise and disgust) as an initial set of face images. These set of input face images are known as the initial set of face images as the training set.

3.1 Feature Extraction using Principle Component Analysis (PCA)

Feature extraction is the process of defining a set of features, or image characteristics, which will most efficiently or meaningfully represent the information that is important for analysis and classification The work in [20] clarified that the magnitude of face images can be reconstructed by the weighted sums of the small collection of characteristic feature or eigenpictures and an efficient way to learn and recognize faces could be to build up the characteristic features by experience over feature weights needed to approximated reconstruct them with the weights associated with known individuals [20]. Each individual, therefore would be characterised by the small set of features or eigenpicture weights needed to describe and reconstruct them, which is an extremely compact representation of the images when compared to themselves [21].

In order to get a tolerable response time, the data dimension needs to be reduced when extracting relevant features from the facial regions. Thus, in this study, Principle Component Analysis (PCA) is used for transforming original images from the training set into a corresponding eigenface. The idea of using principal components to represent human faces was developed by [22] and used by [20] for face detection and recognition. Eigenfaces are a set of eigenvectors used in the computer vision problem of human face recognition. Since eigenface only consists of certain features of the face which may or may not be present in the original image [23]. Hence, the original face image of a person can be reconstructed from if one adds up all the eigenfaces features at the right proportion. Whilst representing an image in a lower dimension with preserving as much information as possible in the embedding space, PCA is a technique that is useful for the compression and robust in treatment of face images with varied facial expression and direction due to the ability of PCA that takes multiple face images as input.

To recognise a face using eigenfaces, initialisation process and recognition process are the two stages involved. Initialisation process includes the following operations:

i. Acquiring the initial set of face images called as training set. As mentioned earlier, this was done in the enrolment process.
ii. Calculate the eigenfaces from the training set, keeping only the *M* images that correspond to the highest eigenvalues. These *M* images define the face space. As new faces are experienced, the eigenfaces can be updated or recalculated.
iii. Calculate the corresponding distribution in *M*-dimensional weight space for each known individual, by projecting their face images on the "face space".

As mentioned earlier, to calculate the eigenfaces from the training set, assume a face image $F(x, y)$ be a two-dimensional N by N array of (8-bit) intensity values. An image is considered as a vector of dimension N^2. Suppose we have M face images, therefore they can be represented as $F_1, F_2, F_{3,}\dots F_M$. The average face of the set is defined by:

$$\psi = \frac{1}{M}\sum_{n-1}^{M} F_n \qquad (1)$$

The mean-centered images can be obtained by subtracting the mean image from each image vector as follows:

$$\Phi_i = F_i - \psi \qquad (2)$$

An example of training set is shown in Figure 3. Each face differs from the average by the vector in (2). Due to this set of very large vectors, it is then subject to principal component analysis, which seeks a set of M orthonormal vectors U_n that adequately describes the distribution of the data. The kth vector U_k is chosen such that:

$$\lambda_k = \frac{1}{M}\sum_{n=1}^{M}(U_k^T\ \Phi_n\ )2 \qquad (3)$$

is a maximum, subject to

$$U_l^T U_k = \partial_{lk} = \begin{cases}1, & \text{if } l=k \\ 0 & \text{otherwise}\end{cases} \qquad (4)$$

The vectors U_k and scalar λ_k are the eigenvectors and eigenvalues, respectively of the covariance matrix.

$$\text{C} = \frac{1}{M}\ \sum_{n-1}^{M}\Phi_n\ \Phi_n^T\ \ = AA^{\ T} \qquad (5)$$

where the matrix $A = [\Phi_1, \Phi_2, \Phi_{3,........}\ \Phi_M]$. The matrix C, however, is N^2 by N^2 and determining the N^2 eigenvectors and eigenvalues is an intractable task for typical image sizes. Therefore, a computationally feasible method is required to find these eigenvectors. If the number of data points in the image space is less than the dimension of the space ($M<N^2$), there will be only *M-1* meaningful eigenvectors, rather than N^2. Meaningful eigenvectors refer to the remaining eigenvectors which will have associated eigenvalues of zero. Consider the eigenvectors v_i of A^TA such that

$$A^TAv_i = \mu_i v_i \qquad (6)$$

Pre-multiplaying both sides by *A*, it can be expressed as

$$AA^TAv_i = \mu_i Av_i \qquad (7)$$

As computed in (7), Av_i are the eigenvectors of $C = AA^T$.

Subsequently, from the analysis, we construct the *M* by *M* matrix $L = AA^T$, where $L_{mxn} = \Phi_m^T\Phi_n$, and identify the *M* eigenvectors, V_l of *L*. These vectors determine linear combinations of *M* training set face images to form the eigenfaces u_l.

$$u_l = \sum_{k=1}^{M} v_{lk}\,\Phi_k, \qquad l = 1, \ldots, M \qquad (8)$$

3.2 Face Classification

Once the eigenfaces are created, identification becomes a pattern recognition task. As described earlier in Section 3.1, the initialisation operations can be performed from time to time whenever there is a free excess computational capacity. The following processes are then used to recognise new face images [20]:

i. Calculate a set of weights based on the input image and the M eigenfaces by projecting the input image onto each of the eigenfaces.
ii. Determine if the image is a face at all (known or unknown) by checking to see if the image is sufficiently close to a "free space".
iii. If it is a face, then classify the weight pattern as either a known individual or as unknown.
iv. Update the eigenfaces or weights as either a known or unknown.
v. If the same unknown person face is seen several times, then calculate the characteristic weight pattern and incorporate into known faces.

Turk and Pentland et al. and Lata et al. clarified that step (iv) and (v) are not usually a requirement of every system. Therefore the steps are left optional and can be implemented as when the there is a requirement [21].

4. FACIAL EXPRESSION DETECTION and CLASSIFICATION

The second authentication in our proposed work is the facial expression detection to verify the particular expression is owned by the genuine user that has passed from the first authentication in Section 3. To date, facial expression is one of the most powerful, natural, and immediate means for human beings to communicate their emotions and intentions. The face can express emotion sooner than people verbalise or even realise their feelings [24].

Existing literatures [25, 26] reported that human facial behaviours could also be useful as another behavioural traits for human identification. In the domain of security, Al-modwahi et al. [13] improved the current surveillance systems by adding facial expression recognition to make a system that detects a person's expression who may intend of causing harm and report to the securities before the person can commit any prohibited work. In 2012, Butalia et al. comprehended that the application of face expression in surveillance and security such as lie detection amongst criminal suspects during interrogation. It is proven that facial cues more often than not can give away a lie to the trained eye [27]. The work in [28] revealed that facial expression is used as the password for authentication to create more randomness and password strength.

However, recognizing facial expression with a high accuracy remains a challenge due to the subtlety, complexity and variability of facial expressions. . To address this problem, we proposed to employ facial landmark detection based on the advantages discussed in [29] [30]

4.1 Facial Landmark Detection

Facial landmark detection is essential in facial expression analysis to find the accurate positions of the facial feature points. Prominent points of a face are usually located on the corners, tips or mid points of the facial components [29]. Tie et al. has mentioned that reliable facial landmarks and their associated detection and tracking algorithms can be useful for representing the important visual features for face registration and expression recognition [29].

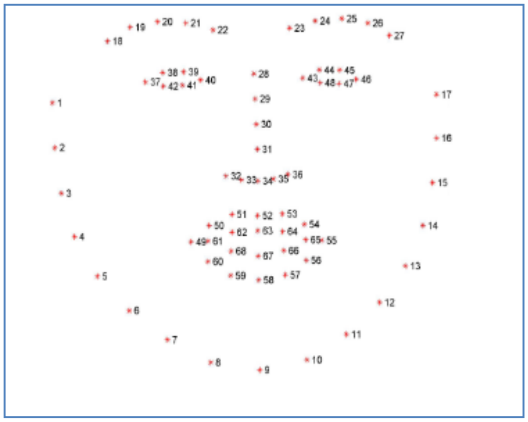
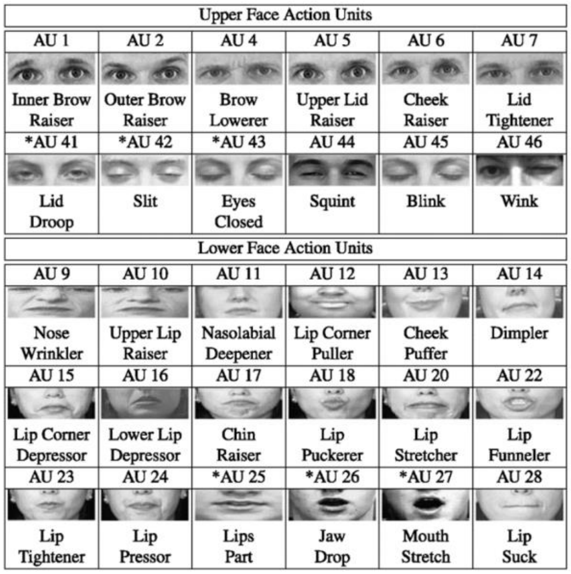


Figure 3: The 68 points mark-up used for our annotations [32].

In order to detect the face expression, we have employed the face landmark localisation approach as mentioned in [31] [30]**.** The points shown in Figure 3 can be concatenated to represent a shape X= $(x_1, \ldots, x_N, y_1, \ldots, y_n)^T$ where (x_i, y_i) denotes the location of the *i*-th point and N is the number of points (N is 68 in Figure 3). Given a sufficiently large number of manually labelled points and corresponding images as the training data. The target of the facial feature point detection is to localise the shape of an input testing image according to the facial appearance. In the training phase, a model is learned from the appearance variations to the shape variations whilst in the testing phase, the learned model is applied to an input testing image to localise facial feature points which refers to the facial shape.

4.2 Action Unit

With the facial feature points have been tracked on the face region, we used action unit from Facial Action Coding System (FACS) to recognise each of the expression. With FACS, observers can manually code all possible discrete movements of the face can be manually coded, which are referred to as action units (AUs). A comprehensive description facial recognition scheme would ideally contain all possible perceptible changes that may occur in a face. This is the goal of a standard FACS which was developed by Ekman and Friesen [33] and has been considered as an empirical study for describing facial expressions. With FACS, observers can manually code all possible discrete movements of the face, which are referred to as action units (AUs). FACS uses forty-four action units (AUs) for the description of facial actions with regard to their location as well as their intensity, the latter either with three or five levels of magnitude. Individual expressions may be modelled by single action units or action unit combinations. Friesen and Ekman introduced such a dictionary for the FACS framework. Ekman et al. [33] presented also a database called Facial Action Coding System Affect Interpretation Database (FACSAID), which allows to translate emotion related FACS scores into affective meanings. Figure 4 portrays partial of the face action unit from the FACSAID. By utilising a decision making system to convert low level face geometry into high level face actions, and then finally into highest level weighted emotion labels [34] such as joy, sad, disgust etc.



Upper Face Action Units					
AU 1	AU 2	AU 4	AU 5	AU 6	AU 7
Inner Brow Raiser	Outer Brow Raiser	Brow Lowerer	Upper Lid Raiser	Cheek Raiser	Lid Tightener
*AU 41	*AU 42	*AU 43	AU 44	AU 45	AU 46
Lid Droop	Slit	Eyes Closed	Squint	Blink	Wink
Lower Face Action Units					
AU 9	AU 10	AU 11	AU 12	AU 13	AU 14
Nose Wrinkler	Upper Lip Raiser	Nasolabial Deepener	Lip Corner Puller	Cheek Puffer	Dimpler
AU 15	AU 16	AU 17	AU 18	AU 20	AU 22
Lip Corner Depressor	Lower Lip Depressor	Chin Raiser	Lip Puckerer	Lip Stretcher	Lip Funneler
AU 23	AU 24	*AU 25	*AU 26	*AU 27	AU 28
Lip Tightener	Lip Pressor	Lips Part	Jaw Drop	Mouth Stretch	Lip Suck

Figure 4: Facial Action Coding System (FACS)

5. OUR PROPOSED WORK

In this study, a fusion authentication model of face recognition and face expression detection is proposed. At the enrolment stage, first time user is required to provide personal details and 16 different poses of face images by choosing one of the seven universal face expression (i.e. neutral, angry, happy, surprise, disgust, fear and sad) as the user's password. All the face images of a genuine user will be stored in the database. Figure 5 illustrates the system architecture diagram of the proposed work.

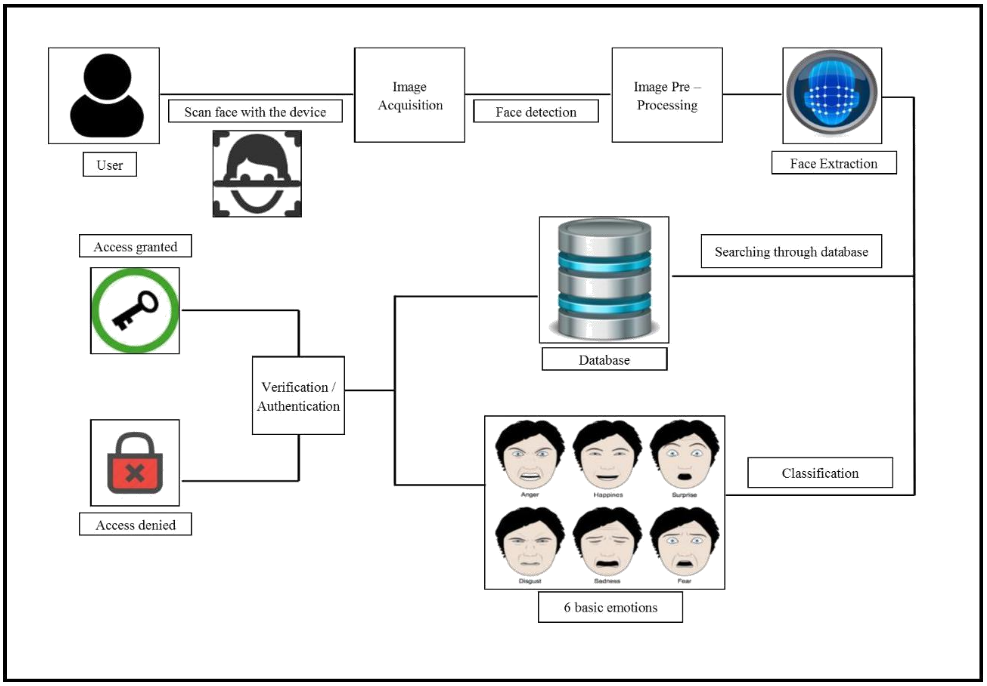


Figure 5. System Architecture Diagram of the Proposed Work

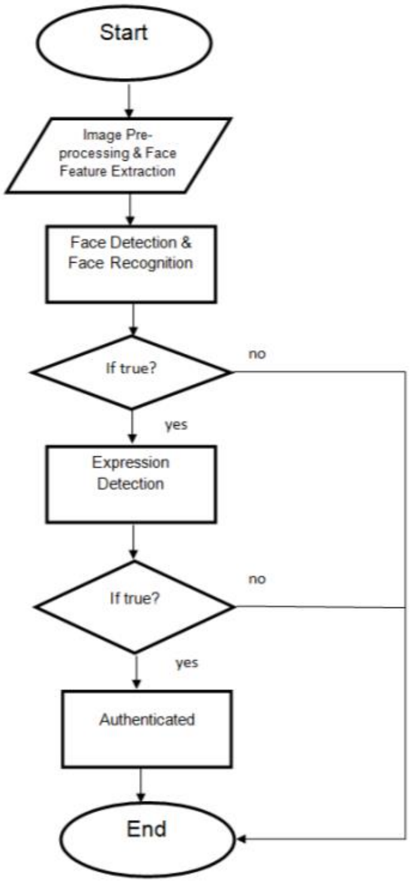

Start
Image Pre-processing & Face Feature Extraction
Face Detection & Face Recognition
If true?
no
yes
Expression Detection
If true?
no
yes
Authenticated
End

Figure 6. Flow Chart of the Proposed Work

During the authentication stage in a controlled environment as depicted in Figure 5, a user is required to scan his or her face via a camera. After acquiring the face image, the system begins to detect the user face before enhancing its image quality in the pre-processing process. For pre-processing, the image frequently takes the form of signal condition (noise removal, pixel position or brightness) with location, segmentation or tracking of the face. To counter the effect of the unwanted transformation such as scaling and rotation of a head in an image, the facial image may be geometrically standardized prior to classification. This normalisation is usually based on reference point of the eyes or nostrils. At the feature extraction process, relevant features from the facial regions are extracted before make comparison in the database. Once the face image is classified as a "known" person followed by the correct face expression which is specifically owned by the user, the identity of the aforementioned individual is verified as an authorised user. Figure 6 illustrates the flow of the proposed work.

6. EXPERIMENTAL TESTING AND PRELIMINARY RESULTS

In order to validate the multimodal biometric-based authentication as proposed in this paper, we have performed a preliminary experiment on the proposed fusion method against single modal biometrics using Japanese Female Facial Expression (JAFFE) as our image database. This dataset is used as the benchmark database for the research in [35] [36]. The database contains ten Japanese females. There are seven different facial expressions, such as neutral, happy, angry, disgust, fear, sad and surprise.

Generally, to measure the performance of a biometric system, many genuine and impostor attempts are made with the system and all similarity scores are saved. By applying a varying score threshold to the similarity scores, pairs of False Rejection Rate (FRR) and False Acceptance Rate (FAR) can be calculated. According to [37], FRR is referred to the proportion of genuine scores that are less than the threshold η, whilst FAR is defined as the fraction of impostor scores that are greater than or equal to η. In another words, FRR is defined as a percentage of impostors accepted by the biometric system whereas FAR is defined as a percentage of genuine users rejected by the biometric system.

6.1 Validation of fusion method against single modal biometric experiment

In the proposed work, we used the 10 subjects (users) from JAFFE database, each subject has 10 samples for recognition and 10 training images as test images respectively. To generate a genuine match score, an image sample from the same user will be compared with the test image from the training set using matcher. Therefore, the total genuine score is 1000 whereas the total impostor score is 9000.

Table 1. The Performance Evaluation of using single modal biometric – face recognition.

Subjects (s)	Genuine Match Score	Genuine Accepted (GA)	Genuine Rejected (GR)	Impostor Accepted (IA)
s1	80	80	20	0
s2	100	100	0	0
s3	100	100	0	30
s4	100	100	0	0
s5	100	100	0	0
s6	90	90	10	10
s7	100	100	0	10
s8	100	100	0	0
s9	60	60	40	0
s10	100	100	0	0

Table 1 shows the performance evaluation results of 10 subjects using purely face recognition. Out of 10 subjects, 7 subjects achieved 100 genuine match score.

In order to get the FAR, we sum up the total of IA from Table 1 over the total of all impostor. Therefore, the FAR is 0.55%. On the other hand, FRR can be obtained by getting the total genuine users rejected over the total of all genuine users. Hence, the FRR of this experiment is 7%.

Apparently. as we can observe the results in Table 1, there are 3 subjects achieved 80, 90 and 60 which is a less promising results in security level. In this scenario, there is a possibility of falsely accepted two or more users that coincidentally have similar facial appearance when using a system based on solely face recognition. Although Table 1 has depicted a relatively high match score achieved when matching one against the other as mentioned earlier. However, if the same system also included facial expression as a unique password selected by the genuine user, pragmatically it would be very unlikely that any given two people would have similar faces and similar selected facial expression.

Eventually, second authentication is essentially needed to overcome the security discrepancies of using purely face recognition. By proposing a fusion biometric authentication, a user can choose his/her preferred facial expression (as shown in Figure 7), either neutral, happy, sad, anger, disgust, fear or surprise as a password. Apart from verifying the identity of a genuine user, the number of impostor users also can be deterred.

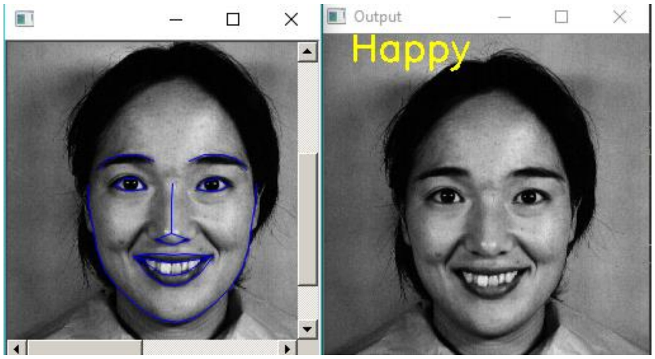


Figure 7. User selection of face expression as a password in second authentication process.

7. CONCLUSION

This paper describes a novel hybrid approach of fusing face and facial expression as a biometric-based authentication model. This research is motivated by utilising the interaction between machine and human facial features in an authentication process. As compared with single modal biometric authentication, this study aims to create a relatively robust, accurate recognition rate and cost effective model. From the performance evaluation of using solely face recognition, facial expression is proposed as a fusion method to improve the genuine match score in an authentication process.

8. ACKNOWLEDGMENTS

Our thanks to Universiti Kuala Lumpur for supporting our research. Besides, we would like to express our gratitude to i-DNA Biotechnology (M) Shd Bhd for giving us a trial version of Face Reader 6.0 to review and study.

9. REFERENCE

[1] M. Kyrnin. 2016. Biometrics will not replace the password, *Tech for Anyone*, 1 September, 2016.

[2] C. L. Deepika, and A. Kandaswamy. 2009. An algorithm for improved accuracy in unimodal biometric systems through fusion of multiple feature sets, *ICGST-GVIP Journal,* vol. 9, no. 3, pp. 33-40.

[3] A. Ross. 2004. Multimodal biometrics: An overview. In *IEEE 12th European Signal Processing Conference,* pp. 1221-1224.

[4] A. Hadid. 2014. Face biometrics under spoofing attacks: Vulnerabilities, countermeasures, open issues, and research directions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* pp. 113-118.

[5] D. B. Mei Yin, M. I. Kamal, N. S. Azmanuddin, S. H. S. Ali, A. T. Othman, and R. Z. W. Chik, Electronic door access control using MyAccess two-factor authentication scheme featuring near-field communication and eigenface-based face recognition using principal component analysis. In *Proceedings of the 10th International Conference on Ubiquitous Information Management and Communication* , pp.1-8.

[6] J. Galbally, J. Fierrez, J. Ortega-Garcia, C. McCool, and S. Marcel. 2009. Hill-climbing attack to an Eigenface-based face verification system. In *2009 First IEEE International Conference on Biometrics, Identity and Security*, pp. 1-6.

[7] H. Ma, O. P. Popoola, and S. Sun. 2015. Research of dual-modal decision level fusion for fingerprint and finger vein image. *International Journal of Biometrics,* vol. 7, no. 3, pp. 271-285.

[8] J. Yang, and X. Zhang. 2012. Feature-level fusion of fingerprint and finger-vein for personal identification. *Pattern Recognition Letters,* vol. 33, no. 5, pp. 623-628.

[9] J. Peng, A. A. A. El-Latif, Q. Li, and X. Niu. 2014. Multimodal biometric authentication based on score level fusion of finger biometrics. *Optik-International Journal for Light and Electron Optics,* vol. 125, no. 23, pp. 6891-6897.

[10] Y. Wang, T. Tan, and A. K. Jain. 2003. Combining face and iris biometrics for identity verification. In *International Conference on Audio-and Video-Based Biometric Person Authentication*, pp. 805-813.

[11] G. E. Determan, V. C. Jacobson, J. Jelinek, T. Phinney, R. M. Hamza, T. Ahrens, G. A. Kilgore, and R. P. Whillock. 2016. Combined face and iris recognition system. U.S. Patent No. 8,705,808. Washington, DC: U.S. Patent and Trademark Office.

[12] J. Neves, F. Narducci, S. Barra, and H. Proença. 2016. Biometric recognition in surveillance scenarios: a survey," *Artificial Intelligence Review*, pp. 1-27.

[13] A. A. M. Al-modwahi, O. Sebetela, L. N. Batleng, B. Parhizkar, and A. H. Lashkari. 2012. Facial expression recognition intelligent security system for real time surveillance. In *Proceedings of the International Conference on Computer Graphics and Virtual Reality* (CGVR) (p. 1). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing

[14] Y. L. Tian, and R. M. Bolle. 2003. Automatic detecting neutral face for face authentication and facial expression analysis.," *AAAI-03 Spring Symposium on Intelligent Multimedia Knowledge Management,* vol. 3, pp. 24-26.

[15] C. Busso, Z. Deng, S. Yildirim, M. Bulut, C. M. Lee, A. Kazemzadeh, and S. Narayanan. 2004. Analysis of emotion recognition using facial expressions, speech and multimodal information. In *Proceedings of the 6th international conference on Multimodal interfaces*, pp. 205-211.

[16] P. I. Wilson, and J. Fernandez. 2006. Facial feature detection using Haar classifiers. *Journal of Computing Sciences in Colleges,* vol. 21, no. 4, pp. 127-133.

[17] P. Viola, and M. J. Jones. 2004. Robust real-time face detection. *International Journal of Computer Vision,* vol. 57, no. 2, pp. 137-154.

[18] P. Viola, and M. Jones. 2001. Rapid object detection using boosted cascade of simple features. *In Computer Vision and Pattern Recognition, Proceedings of the 2001 IEEE Computer Society Conference.* pp. 1-511.

[19] R. Lienhart, and J. Maydt. 2002. An extended set of Haar-like features for rapid object detection. In *Proceedings International Conference on Image Processing* pp. 900-903.

[20] M. Turk, and A. Pentland. 1991. Eigenfaces for Face Detection / Recognition, *Recognition, E., & Neuroscience, C.,* vol. 3, no. 1, pp. 1-11.

[21] Y. V. Lata, C. K. B. Tungathurthi, H. R. M. Rao, A. Govardhan, and L. P. Reddy. 2009. Facial Recognition using Eigenfaces by PCA. *International Journal of Recent Trends in Engineering,* vol. 1, no. 1.

[22] L. Sirovich, and M. Kirby. 1987. Low-dimensional procedure for the characterization of human faces.," *Journal of the Optical Society of America,* vol. 4, pp. 519-524.

[23] V. Hiremath, and A. Mayakar. 2009. Face recognition using Eigenface approach. In *IDT workshop on interesting results in computer science and engineering*.

[24] Y. I. Tian, T. Kanade, and J. F. Cohn. 2001. Recognizing action units for facial expression analysis. In *IEEE Transactions on pattern analysis and machine intelligence,* vol. 23, no. 2, pp. 97-115.

[25] P. Tsai, T. Hintz, and T. Jan. 2007. Facial behavior as behavior biometric? An empirical study. In *2007 IEEE International Conference on Systems, Man and Cybernetics*. pp. 3917-3922.

[26] A. S. Dhavalikar, and R. K. Kulkarni. 2014. Face detection and facial expression recognition system. In *Electronics and Communication Systems (ICECS)*, pp. 1-7.

[27] M. A. Butalia, M. Ingle, and P. Kulkarni. 2012. Facial expression recognition for security. *International Journal of Modern Engineering Research (IJMER),* vol. 2, pp. 1449-1453.

[28] B. Kaliyaperumal, and M. Rajasekaran. 2014. Application Authentication: Facial Expression Password.," *International Conference Image Processing, Computers and Industrial Engineering.* , pp. 135-138.

[29] Y. Tie, and L. Guan, Automatic landmark point detection and tracking for human facial expressions, *EURASIP Journal on Image and Video Processing,* vol. 1, no. 1, pp. 1-15, 2013.

[30] Z. Zhang, P. Luo, C. C. Loy, and X. Tang. 2014. Facial landmark detection by deep multi-task learning. In *European Conference on Computer Vision Springer International Publishing*, pp. 94-108.

[31] E. Zhou, H. Fan, Z. Cao, Y. Jiang, and Q. Yin. 2013. Extensive facial landmark localization with coarse-to-fine convolutional network cascade, pp. 386-391.

[32] C. Sagonas, E. Antonakos, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic. 2016. 300 faces in-the-wild challenge: Database and results. In *Image and Vision Computing,* vol. 47, pp. 3-18.

[33] P. Ekman, and W. V. Friesen. 1977. Facial actio`n coding system. Oxford University Press, USA.

[34] M. Pantic, and L. J. Rothkrantz. 1999. An expert system for multiple emotional classification of facial expressions. In *Proceedings 11th IEEE International Conference on Tools with Artificial Intelligence*, pp. 113-120.

[35] T. Ane, and M. F. K. Patwary. 2016. Performance Analysis of Similarity Coefficient Feature Vector on Facial Expression Recognition. In *Procedia Engineering*, vol. 144, pp. 444-451.

[36] F. Y. Shih, C.-F. Chuang, and P. S. Wang. 2008. Performance comparisons of facial expression recognition in JAFFE database. *International Journal of Pattern Recognition and Artificial Intelligence,* vol. 22, no. 03, pp. 445-459.

[37] A. Jain, A. A. Ross, and K. Nandakumar, *Introduction to biometrics*: Springer Science & Business Media, 2011.