Bi-Modal Methods: An Overview

Olatunji K. A1, Oguntimilehin A2, Adetunmbi3

1 Afe Babalola University Ado Ekiti, Ekiti State, odekunleKenny@yahoo.com
2 Afe Babalola University Ado Ekiti, Ekiti State, ebenabiodun2@yahoo.com
3 Federal University of Technology Akure, Ondo state, aoadetunmbi@futa.edu.ng

Abstract – Various security challenges such as Boko Haram, theft, kidnapping, ISIL, abduction, and so on have been on a high rise as one of the major menace facing our society today. In order to overcome these challenges there is need for identification of the culprits to bring them to book. Unimodal biometric is not enough to combat these security challenges because of its shortcomings which include: spoof attack, noise in the sensed data, inter class variation and so on. Combining two or more biometric features (bi-modal) has been proved to provide better performance than unimodal biometric approach for authentication and verification. This paper presents some literature on biometrics systems that can be employed in achieving a better accuracy in authentication and verification of biometric features. Different kind of fusion strategies to combine these characteristics, different available classifiers and fusion methodologies to achieve greater and accurate recognition performance were also discussed. It is hopeful that researchers in the area of biometrics will find this work very useful.

Key Words – Bi-modal, Biometric, Classifiers, Fusion, Uni-modal.

1 Introduction

Uni-modal biometrics relies on the evidence of a single source of information for authentication (e.g., single fingerprint or face). These systems have to contend with a variety of problems (Arun and Anil, 2004) such as: (a) Noise in sensed data: A fingerprint image with a scar or a voice sample distorted by cold is examples of noisy data. Noisy data could also result from defective or improperly maintained sensors (e.g., accumulation of dirt on a fingerprint sensor) or unfavorable 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), or when the characteristics of a sensor are modified during authentication (e.g., optical versus solid-state fingerprint sensors). (c) Inter-class similarities: In a biometric system comprising of a large number of users, there may be inter-class similarities (overlap) in the feature space of multiple users (Arun and Anil, 2004). Golfarelli et al. (1997) stated that the number of distinguishable patterns in two of the most commonly used representations of hand geometry and face are only of the order of 105 and 103, respectively. (d) Non-universality: The biometric system may not be able to acquire meaningful biometric data from a subset of users. A fingerprint biometric system, for example, may extract incorrect minutiae features from the fingerprints of certain individuals, due to the poor quality
of the ridges. (e) Spoof attacks: This type of attack is especially relevant when behavioral traits such as signature or voice are used.

The above limitations can be overcome by including multiple sources of information for establishing identity (Ross and Jain, 2003) known as multimodal or bi-modal verification system. By using multiple biometric traits, a much higher accuracy can be achieved and biometric systems are expected to be more reliable due to the presence of multiple, (fairly) independent pieces of evidence (Kuncheva et al, 2000).

In this paper some bi-modal characteristics will be examined, methods used in fusion these features and different classifiers will be reviewed.

2 Reviewed Work

Makinde et al (2014) proposed the fusion of face and fingerprint feature extraction using Gabor filter and Mahalanobis Distance Algorithm. An image enhancement technique (histogram equalization) was used to enhance the face and fingerprint images. Salient features of the face and fingerprint were extracted using the Gabor filter technique. A dimensionality reduction technique was carried out on both images extracted features using a principal component analysis technique. A feature level fusion algorithm (Mahalanobis distance technique) was used to combine each uni-modal feature together.

Nayak and Narayan (2013) researched on a multimodal biometric face and fingerprint recognition using neural network system based on adaptive principal component analysis and multilayer perception. In their research, facial and fingerprint images were captured by a digital camera or scanning. The facial images were saved into various formats such as Bitmap, JPEG, GIF and TIFF which later serve as an input into the system. These images were then combined to form a composite biometric trait and process. Signal coming from different biometric channels were first pre-processed, and feature vectors were extracted separately, using specific algorithm and these vectors were combined to form a composite feature vector. There was removal of noise and variation of intensity recorded, sharpening, improving the contrast and stringing the texture of the image. There was also image restoration which extracts image information from a degraded form to make it suitable for subsequent processing and interpretation. Rather than combining the feature vector, they were processed separately and individual matching score was found depending on the accuracy of each biometric matching score which will be used for classification. Each modality was first pre-classified independently. Then the classifier decided whether the image belongs to the face or the non-face class based on the information learned during training. MATLAB was used for the implementation of the research.

Yeong et al (2012) developed a multimodal biometric system using face and irises modalities. In their research images of irises and face were acquired simultaneously using face camera, two iris cameras, near-infrared illuminators and cold mirrors respectively. This device consists of a face camera, two iris cameras, cold mirrors and a near-infrared (NIR) illuminator (including 36 NIR light emitting diodes [LEDs] with a wavelength of 880 nm). They used three universal serial bus (USB) cameras (Webcam C600 made by Logitech Corp) to capture an image containing 1600 by 1200 pixels at a speed of 30 frames/s. The face recognition process was carried out according to the following process. First, the face and eye regions were detected by AdaBoost and rapid eye detection (Viola and Jones, 2004; Kim et.al, 2010). Second, size normalization was conducted to eliminate variations in the detected facial region, while the illumination was normalized using the Retinex algorithm (Hines et.al, 2004). Third, facial features were acquired from the normalized facial image based on principal component analysis (PCA). Finally, the matching score was calculated as the Euclidean distance to provide an input for the SVM. During iris recognition, an iris region was segmented using
integer-based CED and with an eyelid/eyelash detection method (Jeong et al., 2010; Jang et al., 2008; Kang and Park, 2007). Iris codes were generated from the segmented iris region. The matching score of the Hamming distance was calculated and used as the SVM input. These procedures were performed for both the left and right iris images captured using the proposed device. The matching scores for the face and both irises were used as SVM inputs and a final authentication was carried out based on the outputs of the SVM. The SVM required the training database to determine the optimal classifier, so half of the collected images (1,725 images [face images of 575, left iris images of 575 and right iris images of 575, respectively]) were used for training and the remaining ones were used for testing. With the training database, the number of genuine samples (positive samples) was 10,774 and that of imposter samples (negative samples) was 154,251, respectively, in the case of face recognition. In the case of left or right iris recognition, those of positive (10,774) and negative samples (154,251) were also the same as face recognition, respectively. The experimental results showed that the proposed system performs better than face or iris recognition in isolation, as well as the other combination methods.

Jayanta et al. (2012) proposed a biometric authentication where iris and fingerprint features were fused together using score level fusion. The method was based on two components which were quantile transform of the genuine and impostor score distributions and a power transform which further changes the score distribution to help linear classification. After the scores are normalized using the novel quantile power transform.

Sangeetha and Radha (2012) proposed a New Framework for Iris and Fingerprint Recognition Using SVM Classification and Extreme Learning Machine Based on Score Level Fusion. They established that in a Multimodal biometric system, the effective fusion method was necessary for combining information from various single modality systems. Two biometric characteristics were considered in their study: iris and fingerprint. The individual scores of two traits, iris and fingerprint were combined at the matching score level to develop a multimodal biometric authentication system. Two different feature extractors were used for Fingerprint modality and were fused individually with the Iris modality to further evaluate the fusion results. The Centre of Unified Biometrics and Sensors (CUBS) at University of New York at Buffalo developed the first feature extractor utilized for Fingerprint modality. CUBS contain a Chain Code based feature extractor with contour following to detect minutiae as elucidated in (Wang and Han, 2009). Iris Segmentation comprises of two steps: Estimation of iris boundary by using canny edge detection technique and Noise removal by using Hough transform. The iris image was first fed as input to the canny edge detection algorithm that produced the edge map of the iris image for boundary estimation. The exact boundary of pupil and iris was located from the detected edge map using the Hough transform. For Feature Extraction, the normalized 2D form image was disintegrated up into 1D signal, and these signals were made use to convolve with 1D Gabor wavelets. Non-Linear Fusion Methods which includes support vector machine and extreme learning machine were used for classification. K-means Clustering was used to searching the database. In conclusion comparing the classification time, Extreme learning machine performed better than the support vector machine. The experimental results show that comparing SVM and ELM with K-mean cluster methods provide clustering score based on similarity done and reduce the classification time. In this, the Fingerprint-Iris system provides better performance, and comparison of support vector machine and extreme learning machine based on score-level fusion methods was obtained.

Hossain et al. (2012) investigated the role of multi view gait images acquired from multiple cameras - infrared and normal visible images in ascertaining identity. The benefits achieved with multimodal fusion, the roles of efficient subspace features and classifier methods, and the importance of soft/secondary biometric (walking style) in enhancing the accuracy and robustness of gait based identification systems was also examined. Experimental evaluation of several subspace based gait feature extraction approaches using PCA/LDA and learning classifier methods (NB/MLP/SVM/SMO)
on different datasets from a publicly available gait database CASIA was carried out, which showed significant improvement in recognition accuracies with multimodal fusion of multi-view gait images from visible and infrared cameras acquired from video surveillance scenarios.

Mark (2011) incorporated feature data fusion of fingerprint and keystroke dynamics. A new method was developed for fingerprint features representation while feature selection method (normality statistics) was employed for keystroke dynamics to reduce the variability associated with data from the characteristics. An artificial neural network was used in the classification. The keystroke dynamics investigation returned an average FAR of 0.02766095 and an average FRR of 0.0862, which were at least comparable with other research in the field.

Johnson et al (2011) compared Quality-Based Fusion of Face and Iris Biometrics. Q-FIRE database which is a multimodal database composed of face and iris biometrics captured at defined quality levels, controlled at acquisition was used. For the generation of face match scores the FaceIT SDK was used. Iris match scores were then generated using a modified version of the Masek open source software. The fixed sum rule was implemented for score level fusion. Given the K component score vector \( x = [x_1, x_2, \ldots, x_K] \), corresponding to the K matchers, the fused sum of scores \( s_f(x) \) is determined as

\[
\sum_{k=1}^{K} x_k. \quad \ldots (1)
\]

It is important that the scores of both modalities be normalized to a specific range to avoid, as much as possible, a bias toward one of the modalities. In this case, all scores are pre-normalized to the range, using min-max normalization. The second fusion method implemented was the likelihood ratio statistic modeled with Gaussian mixture models (GMM), using an independent training set of match scores. The likelihood ratio statistic requires a training step, which consists of the density estimation of the genuine and impostor distributions, denoted by \( f_{\text{gen}}(x) \) and \( f_{\text{imp}}(x) \). These densities are modeled with Gaussian mixture models (GMM), using an independent training set of match scores, where no subjects included in the training set were used to construct the match score set for evaluation. The likelihood ratio statistic is then defined as the ratio of estimated densities at a given score vector \( x \), as

\[
LR(x) = \frac{f_{\text{gen}}(x)}{f_{\text{imp}}(x)} \quad \ldots (2)
\]

For the third fusion method, the quality measures are exploited in a quality-based likelihood ratio method. This method modifies (2) incorporate quality measure \( q \) and give the following quality-based likelihood ratio statistic:

\[
QLR(x, q) = \frac{f_{\text{gen}}(x,q)}{f_{\text{imp}}(x,q)} \quad \ldots (3)
\]

Kostas et al (2007) described a multi-cue, dynamic approach to detect emotion in naturalistic video sequences. In the quest to receive feedback from the users in an unobtrusive manner, the combination of facial and hand gestures with prosody information was employed to infer the users’ emotional state, relying on the best performing modality in cases where one modality suffers from noise or bad sensing conditions. Recognition was performed via a Simple Recurrent Network which lends itself well to modelling dynamic events in both user’s facial expressions and speech. Nonparametric discriminant analysis with a Support Vector Machine (SVM) which classifies face and non-face areas reducing the training problem dimension to a fraction of the original with negligible loss of classification performance was used for face detection.

Souheil (2007) presented a Multi-modal data fusion for personal authentication using SVM. Elastic Graph Matching was used for face identification. For speech identification, the input signal is
segmented into phonemes and LPC-C (Linear Predictive Coefficients-Cepstrum) vectors are computed and a Hidden Markov Model was used to model the digit-user couple. The LPC-C vectors distribution was modeled by Gaussian parameters which were estimated during training for each digit and for each person. Another Hidden Markov Model was computed to represent the impostor model using a different database of a large number of persons: a digit-world model was obtained for each digit. SVM was used as a classifier. The results of the experiments showed that SVM outperformed the Bayesian conciliation fusion algorithm and that it reduced the total error rate by a factor ranging from 2 to 8 (depending on the type of kernel used for the SVM).

Arpita (2006) proposed three novel techniques for extraction of facial features and recognition of faces from frontal and near-frontal face images called sub-band face representation. This method involved the process of selecting suitable sub-bands of a face, and then reconstructing it using Inverse Discrete Wavelet Transform (IDWT), based on certain criteria. Other two proposed face recognition techniques dealt with two subspaces, namely, range space and null space of within-class scatter which constituted the entire face space if combined. The range space held the entire intra-class variations and the null space contained the intra-class commonalities presented across samples containing variations in expression, illumination and pose. Feature fusion based method used Gramm-Schmidt Ortho-normalization and covariance sum method to combine the discriminative information obtained from null space and range space for constructing a dual space. Also a new approach for combining evidences from face and fingerprint classifiers at the decision level were proposed. In this approach, each of the face and fingerprint classifier was separately exploited on the basis of availability of class specific information to enhance combination performance. Results using face and fingerprint databases, showed that the proposed methodology of using class-specific information at classifier's response outperforms the state-of-art fusion techniques.

Kyong (2004) combine face and ear biometric modalities to improve person’s identification. He proposed a new algorithm for 3D face recognition for handling expression variation. The algorithm uses a surface registration-based technique for 3D face recognition. Normalization was done on 2D image to minimize, to the degree possible, the uncontrolled variations that occur during the acquisition process and to maintain the variations in facial appearance between individuals. Finally, median filtering was applied with a 7x7 kernel window to suppress noise. The face region was interpolated into a 130x150 template that masks out the background. This scales the original image so that the pixel distance between the eye centers is 80. Histogram equalization was applied to standardize the intensity distribution, the “face space” was created from a training set of 2D and 3D images. Initially, eigenvectors with largest eigenvalues are dropped in descending order of eigenvalue, and the rank-one recognition rates of each modality was recomputed each time, continuing until a point was reached where the rank-one recognition rate drops. For distance metrics, Mahalanobis cosine distance metric was used during the matching process. The metric level fusion (late fusion) focuses on combining the match distances that we found in the individual spaces. Having distance metrics from two or more different spaces, a rule of how to combine the distances across the different biometrics for each person in the gallery was applied. The ranks were then determined based on the combined distances. Scores from each modality was normalized to be comparable to each other prior to fusion. The weight was applied to scores (metric) as the combination rules were applied. The multi-modal decision was made as follows. First the probe was matched against the gallery in each of the 3 spaces. This gives three sets of N distances, where N is the size of the gallery. A plain sum-of-distances rule would sum the distances for each gallery subject and select the subject with the smallest sum. They used a confidence-weighted variation of the sum-, multiplication- and min-of-distances rule. For each space, a "confidence" was computed using the first three distances as follows:

\[
weight_{ij} = \frac{metric_{ij} - metric_{i1}}{metric_{ij} - metric_{i4}}
\]  

... (4)
for the $i^{th}$ gallery in the $j^{th}$ modality. Metric $k$ is the $k^{th}$ closest distance measure to a gallery from the observed probes. If the difference between the first and second distance metric is large compared to the typical distance, then this confidence value will be large. The confidence values were used as weights in the metric fusion. Images were acquired at the University of Notre Dame between January and May 2003. Evaluation and compares of the performance of approaches to 3D face recognition based on PCA-based and on iterative closest point algorithms was carried out. The proposed method outperforms 3D eigenfaces when 3D face scans were acquired in different times without expression changes and also with expression changes.

Andrew et al (2004) examined a Bimodal Biometric Verification System Fusion Decision Scheme using Nearest Neighborhood Classifiers. K-Nearest Neighborhood (k-NN) based classifiers were adopted in the decision fusion module for the face and speech experts. The fusion decision schemes considered were voting, modified and theoretic evidence of k-NN classifiers based on Dempster-Shafer theory. The performances of these k-NN classifiers were evaluated in both balanced and unbalanced conditions and compared with other classification approaches such as sum rule, voting techniques and Multilayer Perceptron on a bimodal database.

Wang et al (2003) combined iris and facial features for recognition. Two different strategies for fusing iris and face classifiers were employed. They computed either an unweighted or weighted sum and compared the result to a threshold, then treated the matching distances of face and iris classifiers as a two-dimensional feature vector, using a classifier such as Fisher’s discriminant analysis and a neural network with radial basis function (RBFNN) to classify the vector as being genuine or an impostor.

Claude et al (2002) combined auditory and visual speech modalities for recognition. For acoustic features, cepstral features were used. The cepstrum is the discrete cosine transform (DCT) of the logarithm of the short-term spectrum. The cepstrum can be obtained through linear predictive coding (LPC) analysis or Fourier transformation. Gaussian Mixture Models (GMMs) and Hidden Markov Models (HMMs) were used for recognition as parametric models while non parametric models used were Reference-Pattern Models which take the form of a store of reference patterns representing the voice-pattern space. To counter misalignments, arising from change in speaking rate, temporal alignment using dynamic time warping (DTW) was applied during pattern matching. And Connectionist Models which consist of one or several neural networks such as Multilayer Perceptron MLPs, radial basis functions, neural tree networks, Kohonen’s self-organizing maps, and learning vector quantization.

### 3 Different Fusion Methods

In order to join two biometric inputs, a method called ‘fusion’ is used. Fusion refers to the action of combining two separate biometric inputs (modalities) (Anil et.al, 2005). Information fusion encompasses any area which deals with utilizing a combination of different sources of information, either to generate one representational format, or to reach a decision (Paul, 1982). The information sources are:

i. Single Biometry, Multiple Sensors: The same biometric is acquired by different sensors and combined to complete and improve the recognition process (e.g. optical and solid-state fingerprint sensors). The use of multiple sensors can address the problem of noisy sensor data, but all other potential problems associated with uni-modal biometric systems remain.

ii. Single Biometry, Multiple Instances: The same biometric unit is acquired several times by same sensor and combined to complete and improve the recognition process (e.g. multiple face images of a person obtained under different pose/lighting conditions).

iii. Single Biometry, Multiple Units: The same biometric, but different units are acquired and combined to complete and improve the recognition process (e.g. left and right iris images). This is a recognition
system that works on multiple units of the same biometric measurements (e.g. left middle finger followed by a right thumb). iv. Single Biometry, Multiple Representations: The same biometric is acquired once by a single sensor and different approaches of feature extraction and matching are combined to complete and improve the recognition process (e.g. multiple face matchers like PCA and LDA). v. Multiple Biometrics: Different biometrics of the same person are acquired and combined to complete and improve the recognition process (e.g. face, fingerprint and iris). This approach is the only well-used multimodal biometric fusion scenario (Arpita, 2006).

The information fusion is divided into three parts, pre-mapping fusion, midst-mapping fusion, and post-mapping fusion/late fusion. In pre-mapping fusion information can be combined at sensor level or feature level. Sensor-level fusion can be mainly organized in three classes: (1) single sensor-multiple instances, (2) intra-class multiple sensors, and (3) inter-class multiple sensors. Feature-level fusion can be mainly organized in two categories: (1) intra-class and (2) inter-class. Intra-class is again classified into four subcategories: (a) Same sensor-same features, (b) Same sensor-different features, (c) Different sensors-same features, and (d) Different sensors-different features (Robert, 2005 and Anil et al, 2005).

4 Classifiers

Classification algorithms usually involve some learning - supervised, unsupervised or semi-supervised. Semi-supervised method is usually employed where acquisition of new-labeled sample is small, unsupervised learning is the most difficult approach, as there are no labeled examples. Many bi-modal recognition applications include a tagged set of subjects and most systems implement supervised learning methods. Jain et al (2000) presented three concept of building a classifier, they are: probability, similarities, and decision boundaries.

In similarity approach, patterns that are alike will belong to the same class. The initiative is to create a metric that defines similarity and a representation of the same-class samples. Different techniques can be used such as the Euclidean distance, with the 1-NN decision rule that assign pattern to nearest pattern’s class, a k-means clustering algorithm in unsupervised learning operates like 1-NN but assign pattern to the majority of K nearest patterns, Vector Quantization Method assign pattern to the nearest centroid and has different learning methods, Self-Organizing Maps assign pattern to nearest node, then update node pulling them closer to input pattern, template matching where unlabeled samples are compared to stored patterns by assigning sample to the most similar template which must be normalized.

Some classifiers are built using a probabilistic approach. Examples are Bayesian method that assigned pattern to the class with the highest estimated posterior probability, logistic classifier that predicts probability using logistics curve method and Parzen classifier that combined Bayesian classifier with Parzen density estimates.

Decision boundaries approach’s main idea is to minimize a criterion (a measurement of error) between the candidate pattern and the testing patterns. Examples are the Fisher’s Linear Discriminant (often interchangeably used with LDA) and very much connected to PCA. FLD attempts to model the difference between the classes of data, and can be used to minimize the mean square error or the mean absolute error. Neural networks are another example of nonlinear decision boundary method, in which Multilayer perceptron is a very good case. Also Radia Basis Network that causes the optimization of a multilayer perceptron where one layer at least uses Guassian transfer functions. Nonetheless, neural networks can be trained in many different ways, which can lead to diverse classifiers. They can also provide a confidence in classification with the aim of giving an approximation of the posterior
probabilities. Support Vector Machine that maximizes margin between two classes is another example of a decision boundary classifier (Marques, 2010).

In bi-modal biometrics, there is need to combine classifiers for better recognition performance. According to (Tulyakov, 2008), the classifier combination problem can be defined as a problem of finding the combination function accepting N-dimensional score vectors from M-classifiers and outputting N final classification scores.

There can be several reasons to combine classifiers in bi-modal biometrics: (1) the classifiers that are developed each, with a different approach can be combined for better recognition performance. (2) In case there is different training sets, collected in different conditions and representing different features. Each training set could be well suited for a certain classifier. Those classifiers could be joined together. (3) One single training set can show different results when using different classifiers. A combination of classifiers can be used to achieve the best results. (4) Some classifiers differ on their performance depending on certain initializations.

According to Marques (2010), combiners can be grouped in three categories based on their architecture:

(i) Parallel. All classifiers are executed independently. The combiner is then applied.
(ii) Serial. Classifiers run one after another. Each classifier polishes previous results.
(iii) Hierarchical. Classifiers are combined into a tree-like structure.

According to Jain et al (2000) and Tulyakov (2008) the classifiers combination scheme is presented below.

Table 1: Classifiers combination schemes

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Architecture</th>
<th>Trainable</th>
<th>Info – level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voting</td>
<td>Parallel</td>
<td>No</td>
<td>Abstract</td>
</tr>
<tr>
<td>Sum, Mean, Medium</td>
<td>Parallel</td>
<td>No</td>
<td>Confidence</td>
</tr>
<tr>
<td>Product, Min, Max</td>
<td>Parallel</td>
<td>No</td>
<td>Confidence</td>
</tr>
<tr>
<td>Generalized ensemble</td>
<td>Parallel</td>
<td>Yes</td>
<td>Confidence</td>
</tr>
<tr>
<td>Adaptive weighting</td>
<td>Parallel</td>
<td>Yes</td>
<td>Confidence</td>
</tr>
<tr>
<td>Stacking</td>
<td>Parallel</td>
<td>Yes</td>
<td>Confidence</td>
</tr>
<tr>
<td>Borda count</td>
<td>Parallel</td>
<td>Yes</td>
<td>Rank</td>
</tr>
<tr>
<td>Behaviour Knowledge space</td>
<td>Parallel</td>
<td>Yes</td>
<td>Abstract</td>
</tr>
<tr>
<td>Logistic regression</td>
<td>Parallel</td>
<td>Yes</td>
<td>Rank</td>
</tr>
<tr>
<td>Class set reduction</td>
<td>Parallel/cascading</td>
<td>Yes</td>
<td>Rank</td>
</tr>
<tr>
<td>Dempster – shafer rules</td>
<td>Parallel</td>
<td>Yes</td>
<td>Rank</td>
</tr>
<tr>
<td>Fuzzy integrals</td>
<td>Parallel</td>
<td>Yes</td>
<td>Confidence</td>
</tr>
<tr>
<td>Mixture of local Experts</td>
<td>Parallel</td>
<td>Yes</td>
<td>Confidence</td>
</tr>
<tr>
<td>Hierarchical MLE</td>
<td>Hierarchical</td>
<td>Yes</td>
<td>Confidence</td>
</tr>
<tr>
<td>Associative switch</td>
<td>Parallel</td>
<td>Yes</td>
<td>Abstract</td>
</tr>
<tr>
<td>Random subspace</td>
<td>Parallel</td>
<td>Yes</td>
<td>Confidence</td>
</tr>
<tr>
<td>Bagging</td>
<td>Parallel</td>
<td>Yes</td>
<td>Confidence</td>
</tr>
<tr>
<td>Booting</td>
<td>Hierarchical</td>
<td>Yes</td>
<td>Abstract</td>
</tr>
<tr>
<td>Neural tree</td>
<td>Hierarchical</td>
<td>Yes</td>
<td>Confidence</td>
</tr>
</tbody>
</table>
5 Conclusion

At the end of this paper, we were able to establish that bi-modal authentication performs better than the uni-modal which can be employed to tackle various security challenges (like abduction, theft, kidnapping, boko haram, and so on) facing our society today in order to bring the culprits to book. Fusion strategies were also presented to combine more than one feature and classification techniques and their fusion were reviewed to facilitate effective recognition.

References


Emdad Hossain, Girija Chetty and Roland Goecke (2012). Multi-view Multi-modal Gait Based Human Identity Recognition From Surveillance Videos. *University of Canberra, Australia*, emdad.hossain@canberra.edu.au


Kyong I. Chang (2004). New Multi-Biometric Approaches for Improved Person Identification. A Dissertation Submitted to the Graduate School of the University of Notre Dame in Partial Fulfilment of the Requirements for the Degree of Doctor of Philosophy.

Mark Abernethy (2011). User Authentication Incorporating Feature Level Data Fusion of
Multiple Biometric Characteristics. A thesis is presented for the degree of Doctor of Philosophy Murdoch University, January 2011.


