Cascaded multimodal biometric recognition framework

Asim Baig[1,2], Ahmed Bouridane[3], Fatih Kurugollu[1], Badr Albesher[4]

[1]*Institute of Electronics, Communications and Information Technology (ECIT), Queen's University Belfast, BT3 9DT, UK*
[2]*Department of Electrical and Computer Engineering, Muhammad Ali Jinnah University, Islamabad, Pakistan*
[3]*Department of Computer Science and Digital Technologies, Northumbria University, Newcastle Upon Tyne, NE12 1XE, UK*
[4]*Deanship of Scientific Research, Al Imam Mohammad Ibn Saud Islamic University (IMSIU), P.O. Box 37254, Riyadh 11439, KSA*
E-mail: ahmed.bouridane@northumbria.ac.uk

Abstract: A practically viable multi-biometric recognition system should not only be stable, robust and accurate but should also adhere to real-time processing speed and memory constraints. This study proposes a cascaded classifier-based framework for use in biometric recognition systems. The proposed framework utilises a set of weak classifiers to reduce the enrolled users' dataset to a small list of candidate users. This list is then used by a strong classifier set as the final stage of the cascade to formulate the decision. At each stage, the candidate list is generated by a Mahalanobis distance-based match score quality measure. One of the key features of the authors framework is that each classifier in the ensemble can be designed to use a different modality thus providing the advantages of a truly multimodal biometric recognition system. In addition, it is one of the first truly multimodal cascaded classifier-based approaches for biometric recognition. The performance of the proposed system is evaluated both for single and multimodalities to demonstrate the effectiveness of the approach.

---

1 Introduction

Multimodal biometric recognition is defined as a method which utilises two or more biometric modalities together during the recognition process. The advantage provided by the use of a multimodal biometric recognition system is that the fusion of matching scores from two modalities improves overall performance of the recognition system. An additional advantage of a multimodal biometric recognition system is that since all the biometric traits are required as part of the input to the system before recognition takes place it makes spoofing or cheating the system very difficult. If a practical biometric recognition system is to be developed then the advantages outlined above make a multimodal biometric recognition system an attractive solution on paper. It should improve accuracy and population coverage as well as overall security of the system. However, these systems do have some serious drawbacks that make implementing them for a practical application very difficult. The key problems are related to processing time and system's complexity which hinder the development of a practical multimodal biometric recognition system.

One solution is to use simple, low complexity but fast unimodal biometric recognition systems as the building blocks of a multimodal system. However, these low-complexity unimodal systems are usually not very stable. They tend to make the overall system unstable when used in the traditional score level fusion-based multimodal biometric system and may affect overall system performance. This paper proposes a unique framework to leverage these low-complexity systems in developing a very stable and effective multimodal biometric system.

Cascaded classifier-based systems are a type of ensemble-based learning approach. Ensemble-based learning approaches are a branch of machine learning which studies the effects of a collection of classifiers. Cascading-based approaches are based on concatenating several classifiers such that the output of each classifier is utilised as additional information for input of the next classifier in the cascade. The difference between the cascading-based approach when compared with a voting or stacking-based approach is that while the former is a multistage system the latter is considered as a multi-expert system. The two most widely used cascading based approaches are boosting [1, 2] and bootstrap aggregation or bagging [3].

Cascading-based approaches especially boosting approaches are based on the following question posed by l Kearns and Valiant [4]: *can a set of weak learners create a single strong learner?* A weak learner is a classifier that can label examples somewhat better than random guessing, that is, it is only slightly correlated with true classification. On the other hand, a strong learner is arbitrarily well correlated with true classification.

Most cascading-based algorithms consist of a set of weak learners such that each weak learner is trained on a given distribution and added to the final strong learner with a weight that is usually related to the accuracy of the weak learner. These algorithms have been used quite successfully in object detection applications such as face detection [5] and handwriting recognition [6] etc.

It is worth noting that very limited work had been done so far on the use of cascaded classifiers for biometric recognition systems. One of the earliest examples of using cascaded classifiers in a biometric recognition system is described in [7] where the authors utilise two iris recognition classifiers, a local feature-based classifier (LFC) and an iris blob matching-based classifier to improve the performance of an iris recognition system. Their approach is to construct a two stage classifier with reject option. The LFC is implemented as the first stage and the blob matcher stage is consulted when LFC is uncertain regarding its results. More recently, Iqbal and Namboodiri [8] used a cascade of simple linear projections on random lines to filter the database and reduce the number of users for final matching stage. They claim to reduce the search space by 60% without increasing the FAR. In [9], the authors use cascaded classifiers like AdaBoost and RankBoost to perform score level fusion. They claim that Adaboost can provide the best score level fusion as well as increase the area under the curve of receiver operating characteristic (ROC) curve. Similarly, in [10], the authors use three different classifiers for three different features extracted from the Gabor coefficients of a single signature for signature verification. Different sets of neural networks are trained on these features to evaluate them and the results are fused using either majority voting, weighted majority voting or cascaded classifier. In their cascaded classifier-based approach the neural network classifiers operate serially and when a simple classifier accepts a signature, the ensemble classifier accepts it. The authors claim that the cascaded classifier-based approach provides best possible results with minimum false accept rate (FAR).

A review of the research on multimodal biometric recognition systems show that not many attempts have been made to develop cascaded classifier-based approaches. A reason is that there is very little correlation between different types of biometric traits. Therefore it is difficult to develop a framework where the output of one unimodal biometric recognition system can be used as part of the input of the next unimodal system to improve overall performance. Erzin et al. [11] made one of the earliest attempts at developing a cascaded multiple modality biometric system. They attempted to fuse the results from audio, face and lip motion modalities using an adaptive cascade of classifiers. The adaptive cascade of classifiers approach proposed by them was to evaluate each modality separately and if one of the modalities provides a reliably strong accept or reject decision other scores are ignored. If no single modality provides a dominant score the decision is taken according to the classifier with the highest likelihood offset amount the classifiers. Lee et al. [12] proposed a fusion strategy that fuses multiple physical traits in a cascade structure, in which users are verified with individual modules sequentially in separate stages, each stage contains a unimodal module. Once the user is verified with one module the result is accepted and the rest of the modules are not processed. The authors proposed the sequence of face, voice and iris to develop the cascade. Lakshmiprabha et al. [13] proposed a similar approach but instead of using different modalities the authors used different features of the face to develop the serial cascade. The proposed cascade works in the same way in sense that once the user is verified by one module the rest of the modules are not processed. Similarly, Soviany et al. [14] proposed a multi-classifier approach that uses a post-classification biometric fusion method in which the biometric data classifier outputs are combined in order to improve overall biometric system performance by decreasing the classification error rates. They evaluate all the three modalities simultaneously and then use weighted fusion function to provide the final decision.

It is interesting to note that of all the approaches outlined above only the one presented by Iqbal and Namboodiri [8] attempt to filter the database and generate a candidate list to reduce search space in somewhat similar fashion to our approach. Their results show a significant reduction in search space, in fact, they claim a reduction of search space by about 60% without increase in FAR. Although it is a considerable reduction in search space yet it is dependent on selection of optimal values for multiple parameters such as feature representation, projection window width, cascade sequence and is based on a single modality. The cascaded system proposed in this paper, provides a significantly larger reduction in search space with improved FAR and accuracy with fewer parameters to optimise as discussed later in this paper.

The use of multiple modalities in the proposed approach may help in reducing FAR and improving accuracy of the system. The proposed system does not use matching score output of one system as part of the input for the next one, instead it operates by exploiting the quality of matching scores for each stage to reduce search space for the next stage. This proposed framework not only allows for development of a cascaded biometric recognition system utilising two different modalities but also leverages one of the key properties of a cascaded classifier-based system namely to develop a strong classifier by combining a set of weak classifiers together and reducing search space after each classifier in the cascade.

This paper presents a framework based on cascaded classifiers such as boosting and bagging approaches for multimodal biometric representation. Section 2 introduces the basic principles of cascading classifiers and briefly outlines the problems associated with using cascaded classifiers for multimodal biometric recognition. The experimental system is presented in Section 3 and the test datasets used to evaluate the proposed framework are outlined in Section 4. Section 5 comments on how to address the issues outlined in Section 2 and describe the proposed framework in detail. Section 6 presents the experimental results. These results are then evaluated and discussed later in the same section.

2 Cascaded multimodal framework

Cascaded classifiers are considered very useful in multiple machine learning and computer vision applications. This is because they tend to be more stable and easier to implement than a single strong classifier. Although these classifiers require relatively longer time to train, once trained they perform faster than a single strong classifier. These properties make them an attractive option in design and development of a multimodal biometric recognition system.

A simple cascaded classifier-based system designed for object detection is developed such that each weak classifier is trained on a single feature of the object. Therefore if the cascade consists of a set of *n* weak classifiers then *n* different features of the object to be detected need to be extracted. Generally, these features are highly correlated as they are generated from the same object. This correlation between the features and their values is an important factor in proper working of a cascaded classifier-based system.

A cascaded classifier-based system operates in the following way; the template image is input to the first weak classifier and each part of the image is labelled as being part of the required object or not. The portions of the images deemed to be part of the object to be recognised are presented as the input to the next weak classifier. This reduces the search space for the next weak classifier. The process is repeated until the final result is produced by the last classifier.

The idea of cascading multiple weak classifiers to generate a strong classifier works for applications such as object detection but it does not seem to work so effectively for a multimodal biometric recognition system. The reason is that although a multimodal biometric recognition system does consist of multiple classifiers, the input of each of these classifiers is a different biometric trait. Since the biometric traits are different their extracted feature sets are generally not correlated. Therefore there is no evidence to show that the output of one classifier would have any exploitable relationship with input of the next classifier. It should be noted that a cascaded classifier-based approach does work for a single input-based biometric recognition as shown by the author of this paper in [15]. The reason is that since all the classifiers have the same input; it is possible to use the output of the first classifier in the ensemble to enhance the performance of the next classifier thus improving the overall performance of the system.

In a multimodal biometric recognition system, if each classifier in the cascade is able to reduce the number of possible candidates in the enrolled users' database, the resulting reduced candidate list should improve the chances of a correct match provided genuine users remain part of the shortlisted dataset. This reduction in user database can be achieved by analysing the results of each classifier in order to eliminate those enrolled users from the possible candidates list for the next classifier that are confirmed to be impostors. Each stage of the cascaded classifier system should then further reduce the candidate list for matching at subsequent classifier stage thereby providing an overall improved matching result.

The term candidate list in this context implies the list of shortlisted candidates generated from the enrolled users' database. This candidate list is in fact a reduced form of enrolled users' database. Therefore the key to developing an effective cascaded classifier-based approach is to limit the number of enrolled users in the candidate list for use by the next classifier. A match quality measure can be used to reduce the candidate list. The idea is to train the classifiers in the ensemble to provide a list of potential users for the next classifier based on the utilised match quality measure. It should be noted that the quality measure and its related parameters should be selected carefully such that the candidate list generated contains genuine users, and remaining as small as possible.

Fig. 1 shows the block diagram for a two stage cascaded multimodal biometric recognition system that uses two biometric inputs which are processed sequentially. Initially, the first biometric trait is input and the first matching algorithm evaluates it against the complete database. The output match scores are then used to calculate the top-N users in the database so as to generate the corresponding candidate list. This candidate list is then passed to the second biometric trait as an additional parameter and the matching algorithm processes the input against the enrolled users from within this candidate list and the result is output. As mentioned before, the last matcher has to be a strong matcher. Therefore in this case the second matcher will be a strong one and will provide the final result. The reason for using a strong matcher at the end of the ensemble is that strong matchers are expected to provide good results on large datasets and with decrease in the size of the dataset the chances of getting correct results increases. This is one of the key advantages of this framework in that for each classifier the candidate list is reduced thereby reducing error rates resulting in a stable and accurate system.

To generate this top-N candidate list, we need to evaluate the quality of the match performed by the matching algorithm. If the match quality is high then the candidate list will contain fewer users and vice versa. To quantify the quality of matching we propose to use the Mahalanobis distance-based match quality matrix presented in [16], in conjunction with the number of side-lobes as the configuring/tuning parameter. Mahalanobis distance is actually a similarity measure but by using it to calculate the similarity between genuine and impostor distributions we can generate a quality measure as shown in [16]. If similarity between the two distributions is high then the match quality is low, on the other hand, if the similarity is low then the match quality is high. The match quality matrix or score is generated by considering each match score vector to be genuine match score vector and evaluating the Mahalanobis distance [17] between this genuine match score and the rest of the match scores. For an NxM match score matrix the result will be a $Nx1$ matrix containing the Mahalanobis distance between the match score vector for each input and the rest of the match scores. All the classifiers (except the last one) are trained to generate a top-N candidate list depending upon the number of side-lobes present in the match score quality measure. Each consequent classifier performs the matching only on the users of its corresponding candidate list. Unlike the traditional cascaded classifier-based systems, the last classifier in this ensemble is always selected to be a strong matcher. This allows the last classifier to perform matching on a shortlisted candidate list and outputs the match score and matching decision.

The reason for using the Mahalanobis distance-based quality measure as the filter to generate the candidate list is that it is modality independent. This allows for the same filter to work irrespective of the modalities utilised, thus providing flexibility in implementation of the proposed framework. Another advantage of using this quality measure as a filter is that as it is evaluated from the match scores it incorporates effects of input image quality, the feature extractor as well as the matching algorithms. This enables us not only to generate the candidate list after deployment but also to comment on the quality of input images and the effectiveness of the selected matcher for use in the cascade during the development process. A more detailed discussion on these development time benefits is provided a little later in this section.

The statistic of this quality measure selected to quantify it for evaluation and to generate the candidate list is side-lobe count. Side-lobe count was selected because not only is it a simple statistic to calculate, thus saving processing time, it also provides a stable result for the majority of modalities. This allows for the use of a single statistic for all the different modalities used in the system.

The aim of the proposed framework is to not only improve system performance but also to improve its processing speed. Experimental results presented in Section 6 show improvement in performance and processing speed. In terms of improvement in processing speed the rationale is that for a strong matcher utilised in the cascaded classifier system, in standalone mode it will have to process all the enrolled candidates in the database. On the other hand, in a cascaded system it will only process the short listed candidates.

3 Experimental test bed

A suitable test bed is necessary to evaluate the performance of the proposed framework. The first aspect when designing the test bed is to identify the modalities to be used in development of the cascaded classifier. In the limited time scale of the study, it is nearly impossible to test the proposed frameworks on a large collection of modalities and their combinations. Therefore for the sake of maximising time and effort, two modalities are selected for the test bed, namely: fingerprint and iris modalities traits. There are two reasons for using fingerprint modality. Firstly, it is the most commonly used modality therefore there is a very strong chance that in the design of any practical multimodal biometric recognition system fingerprint modality will be present. Thus, testing the frameworks on fingerprint modality provides a preliminary performance analysis for any future practical implementations. Secondly, to properly understand the working of a biometric recognition system, which is critical to the development of a stable framework, it is important to have the experience of implementing at least one biometric identification system. Standard algorithms for a fingerprint based identification system are well documented and can be implemented quite quickly all the while providing a invaluable insight into the true working of a biometric recognition system. These are the two main reasons for selecting the fingerprint as one of the modalities. The choice of iris for the second modality was simply because it is one of the most stable modalities in terms of its performance and as such provides an ideal candidate for the final stage of the system.

One of the most effective methods for demonstrating the usefulness of a framework is to use reasonably weak components in development of the test system and then evaluate it using standard datasets. The rationale is that if the framework performs comparably with weaker components its performance will definitely be better with state of the art components. Therefore the test bed has utilised the following different and relatively weaker components.

3.1 Feature extractors

The following different feature extractors are used in the test bed to perform testing of the proposed frameworks.

For *Fingerprint* data the following feature extractors are used:

- *Feature extractor* 1 *(FE*1*):* A chain code-based feature extraction approach using contour following to detect the minutiae as described in [18]. The code for this feature extractor was provided by the Centre of Unified Biometrics and Sensors (CUBS) at the University of New York at Buffalo. This feature extractor provides a stable set of minutiae points even in noisy input images.

- *Feature extractor* 2 *(FE*2*):* A simple binarisation and thinning-based minutia extractor consisting of a segmentation stage [19], an enhancement stage utilising high-boosting filtering approach, a binarisation stage using Niblack approach [20], an eight-connected minutiae detector and a line tracing approach to remove spurious minutiae [21]. This feature extractors providing spurious minutiae in noisy images but has the same performance as the first feature extractor for good quality images.

It is important to point out here that, as per the second reason given above for the selection of fingerprint modality as part of this test bed, FE2 was implemented. Although the approaches used in the implementation are standard we were able to provide considerable improvement to the segmentation module and the resulting findings were published in [20].

The feature extractor employed for *iris modality* is based on Daugman's approach [22] and it is

- *Iris FE*1 *(IFE*1*):* Implemented by Libor Masek as described in [23].

This feature extractor generates a bit stream from the iris referred to as iris code by Daugman. Hamming distance is then used to provide the matching score.

3.2 Matchers

The details of the two matchers used for testing the proposed frameworks are provided below

- *Matcher* 1 *(M*1*):* A graph-based matching approach underlined in [24] and provided by CUBS is used for matching. This is a strong matcher based on a graph transversal algorithm for local minutiae neighbourhood-based matching. It is used purely for fingerprint feature sets.

- *Matcher* 2 *(M*2*):* A simple Hamming distance-based matcher. It can be used for both fingerprints and iris feature sets. Although it is being used exclusively for iris matching in the proposed test bed.

It is important to point out here that apart from FE2 all other codes were taken from the respective developers and are treated as 'black boxes', that is, no modification was made to them. This was done to ensure that any changes in accuracy or stability of operational conditions within the test systems were solely because of the proposed framework. This was also achieved by using the same algorithms for each and every system designed in the course of the research including the systems based on the conventional approaches or the proposed framework. The systems are then evaluated using the same datasets generated from a standard database as described in the next section. Another important point to note here is that wherever the performance of the systems based on the proposed framework is assessed against the systems presented in various research literatures, the comparison is done in terms of percentage of improvement rather than the actual scores since the algorithms and test datasets are extremely different.

3.3 Test datasets

A truly multimodal biometric database is required to properly evaluate the efficiency of any multimodal biometric recognition system. Therefore, to gauge the effectiveness of the proposed frameworks datasets for each of the selected modality are developed from the multimodal database provided by West Virginia University. The complete West Virginia University Database (WVUdb) contains six different modalities, that is, hand geometry, fingerprint, iris, palm-print, voice and face but the university does not allow access to the complete database. They generally provide a few modalities for download and face modality is always provided separately as a unimodal database. Further information about the database can be found in [25]. To test the proposed framework the fingerprint and iris modalities of the WVUdb were used. Some details about the images acquired for these modalities are provided below.

- *Fingerprint:* SecuGen optical fingerprint scanner was used to acquire the fingerprint images. Uncontrolled image acquisition is performed without cleaning the glass plate of the scanner between acquisitions. The image acquired is of size 292 × 248 (72 kB). Each fingerprint image is coded with numbers from 1 for thumb to 5 for pinkie along with a 'L' for left hand and 'R' for right hand, for example, *R*1 means thumb of right hand. Five images of each finger were acquired with the user instructed to lift the finger between each acquisition.

- *Iris:* OKI IRISPASS-h handheld device was used to acquire the iris image. The user was instructed to hold the device at a fixed distance from the eye and to cover the other eye. The image acquired is of size 480 × 640 (302 kB). Generally four images of each eye were acquired but additional samples were also acquired if the image was deemed to be of poor quality subjectively.

Each user in the database is assigned a random seven digit user identification code at the first acquisition. Owing to the way data acquisition process is designed for the database different modalities contain different number of images for each user. It was therefore required to partition the data in such a way that equal number of enrolment and template images are available for both fingerprint and iris modalities of each user.

To generate an unbiased dataset, five hundred (500) unique fingerprint images of 100 different users were acquired randomly; care was taken to avoid any overlap. This means that each user was assigned 4 enrolment images, which were pre-processed and stored in the users' database and 1 input image. Once the fingerprint images are finalised the users' identification number is used to acquire five (5) iris images of the identified users. This allows for the development of a set of an unbiased, random datasets of fingerprint and iris images for 100 users each with 4 enrolment images in the users database and 1 input image each. For sake of evaluating and explaining the results fingerprint dataset is referred to as WVUfd and iris dataset is referred to as WVUid in the next section.

3.4 Proposed system design

This section outlines the parameters required to design a system based on the proposed framework and how to tune them for optimal performance. The first step is to select a criterion for the generation of the top-*N* candidate list. To determine the number of users in the top-*N* candidate list the number of side-lobes of the match quality score [16] are plotted against the position of genuine matches in an ordered match score vector. The match score vectors are ordered in descending order. Fig. 2 shows the plot of these side-lobes against the genuine match position in the match score vector for a fingerprint modality-based biometric recognition system. The plot is based on the match score form the system developed using algorithms *FE*1 and *M*1. A more detailed look at the algorithms is provided in Section 3.

An analysis of similar plots for different biometric matching algorithms provides some interesting insight. Figs. 3 and 4 depict the number of side-lobes from the quality measure and the positions of the genuine matches based on algorithms *FE*2 and *M*1 and *FE*1 and *M*2, respectively.

As discussed above an important development/design time observation is that random plots are not suitable for use in a cascaded classifier-based framework since the random plot indicates an unstable system where the positions of the genuine scores are not correlated with the number of side-lobes. Fig. 2 shows a more stable system, which can be used in cascaded classifier-based framework. On the other hand, Figs. 3 and 4 show systems that are not deemed suitable.

An additional property of these plots, which was highlighted above, is that they are also able to indentify low-quality input images. The outliers in a plot of a stable system can indicate low-quality images. For example, in Fig. 2 an input image has the genuine match score at 62nd position when the side-lobe count is only 6. This is a clear outlier and it is produced specifically because of the low quality of the input image. The corresponding low-quality images are shown in Fig. 5. It is interesting to note that a visual inspection of these low-quality images seems to show a reasonable quality. The reason that such an image is of low quality is because it has a lot of noise in the middle and that will cause the feature extractor to generate a large number of spurious minutiae. Similarly, any input image with the position of the genuine match score above 30 can be considered low-quality images. The input images with genuine match position less than 30 but more than 5 can be considered medium quality images and the ones with a genuine match position of less than 5 can be considered good quality images. Figs. 6 and 7 show some examples of medium and good quality images. On visual inspection a medium quality image may be identified as being of low quality due to the fact that a large area of the image is totally blacked out because of noise. The reason why this image performs better than the two shown in Fig. 5 is because in most algorithms the black area is ignored by the feature extractor and as such does not contribute any false minutiae points whereas the rest of image is of good quality and generates a set of accurate minutiae points.

Two possible strategies are available for selecting the size of the candidate list once the biometric modality is selected for a certain stage. The first approach is to select a fixed number of users for the candidate list each time. The second approach is to use the results of the position of the genuine matches against the number of side-lobes in quality measure to develop a variable candidate list size strategy. Suppose that the unimodal biometric recognition system uses $FE1$ and $M1$ algorithms and is considered for the first stage of a two stage cascaded classifier-based framework. In this case, the results of Fig. 2 are used to calculate the candidate list size. For the static candidate list size it is obvious that the candidate list will contain top 60 users, whereas, if the system considers the various number of side-lobes generated by an input image then a more dynamic candidate list size can be established based on the criteria shown in Table 1. It should be noted that the outliers with low-quality images are ignored during the calculation of the candidate list. This may induce a negligible error into the whole system. Table 1 shows the possible sizes of candidate lists calculated from Fig. 2.

A quick look at Fig. 2 shows that (barring the outliers) for the input with less than ten side-lobes in the quality measure the highest position of the genuine user in top-*N* candidate list is around 12th. Similarly for inputs with quality measure side-lobes between 10 and 20 the highest position of the genuine user is around 30th and for the rest the highest position of the genuine user is around 55th. Table 1 uses these values as thresholds to calculate the size of the candidate list. It should be noted that these thresholds can be fine tuned depending upon the dataset and type of algorithms used.

The cascaded classifier-based biometric recognition system designed based on the proposed framework will then work in the following way. The input image for the first modality will be passed to the first stage of the cascade. This image will then be compared with all the users in the database. Once the match scores become available the proposed quality measure will be evaluated and the side-lobes counted. The new candidate list for the next stage and the modality will be generated based on the conditions outlined in Table 1.

The above discussion outlines an important design parameter, which is the calculation of the cutoff threshold for the number of users in the candidate list. This cutoff threshold should be evaluated by generating a table similar to Table 1 based on the performance of the selected unimodal biometric system on the provided training data.

The improvement in the processing speed (if any) thus obtained would depend on the size of the candidate lists. If the candidate lists are consistently of smaller size then the improvement in processing speed would be significant. To evaluate this Fig. 8 plots the total count of users against the number of side-lobes. Fig. 8 shows that when using the thresholds outlined in Table 1 over half the inputs to the first stage of the proposed biometric recognition system will generate a candidate list with only 15 users. It is also interesting to note that less than 10% of the inputs would generate a large candidate list (i.e. with over 60 candidates); even then the candidate list is 40% smaller than the complete database. This shows that the number of users being processed by the strong matcher reduce considerably when using the proposed framework thus providing improvement to the overall processing speed of the system.

As pointed out before, Iqbal and Namboodiri [8] also use a somewhat similar idea for a unimodal system. They try to reduce the search space after each stage of the cascade and provide a reduced candidate list for the next stage. They claim the best reduction rate of 62.1% on the dataset used without increase of FAR. In comparison, the framework proposed in this paper provides the best possible reduction of 85% of the datasets used and the worst possible reduction rate of 40%. It is also interesting to note that since the proposed framework uses dynamic thresholding of the candidate list only 10% or less of the inputs provide 40% reduction rate, majority of the inputs provide a reduction far better than claimed by Iqbal and Namboodiri [8].

4 Experimental results

A fair assessment of the improvements provided by the proposed framework is difficult to achieve since there exists no similar system. A comparison between the proposed framework and a standard multimodal fusion-based biometric recognition system is also not reasonable. The reason being that the proposed framework is not a fusion-based approach. A standard multimodal fusion-based system combines the results from the two constituting systems to generate a final matching score vector. This matching score vector (which is influenced by both the matching algorithms) is used to make the decision. On the other hand, the proposed framework simply reduces the size of the candidate list in order to improve the performance of the unimodal biometric recognition system at the end of the cascade. Therefore the only justifiable option is to perform the comparison against a strong unimodal biometric recognition system. The following configuration was used from the set of available algorithms (as documented in Sections 3 and 4). The iris modality-based biometric recognition system used to test the proposed framework was configured using ***IFE1*** as the feature extraction algorithm and ***M2*** as the matcher. The ***WVUid*** database was used for testing. A unimodal biometric recognition system based on fingerprint modality was used to reduce the candidate list. This system consists of FE1 and M1 (as discussed in Section 2) and uses WVUfd as the database. For sake of this experiment, first the iris modality-based biometric recognition system was executed on the complete database (i.e. 100 users and 4 enrolment images each). Later, the fingerprint-based system was used to reduce the candidate list and the iris-based system was applied on the reduced candidate list. The results are shown in the second row of Table 2.

The results provided above are evaluated with equal error rate (ERR) as the operating threshold. The reason *IFE*1 is used in these experiments is that it provides accurate and stable results. An important point to note when developing a biometric recognition system based on the proposed cascaded classifier-based method is that a weaker classifier should be used for sections that reduce the candidate list and a stronger classifier should be used at the end of the ensemble to obtain the final decision. Using weaker classifiers to reduce candidate list not only enables faster execution, it also generates a longer candidate list thus ensuring that the genuine users are found in the candidate list. The final classifier in the ensemble works as a standard unimodal biometric recognition system, therefore it should be as strong as possible to provide the best possible results. Fig. 9 shows the ROC curves for the three systems evaluated in Table 2. It shows that the fingerprint–iris cascade-based system developed on the proposed framework maintains almost similar FAR and false reject rate (FRR) to that of the unimodal iris recognition system.

In order to further evaluate the performance of the proposed framework Figs. 10 and 11 show the results in terms of FAR, FRR and accuracy at different thresholds for the iris recognition system and the fingerprint–iris system. It can be clearly seen from the results that the proposed framework improves the accuracy of the system while maintaining the error rates. In fact, the proposed system improves the performance by about 10%. It is worth pointing out that the two (2) imposter matches in the proposed system are due to the two very low-quality partial images that were identified above in Section 2. The candidate list generated by these two images did not contain the genuine user; this is due to the fact that, for these outliers the genuine match position was around 60th whereas based on the threshold the candidate list contained only top 15 candidates. Therefore, using these candidate lists was always going to evaluate to an incorrect match.

To thoroughly evaluate the proposed framework a number of additional questions must be answered. Firstly, what would happen if the algorithms in the cascade were switched? Would the performance remain the same or change? Secondly, how does the proposed framework perform on a single modality? Finally, would the performance of the system keep increasing with the addition of additional short-listing stages? If not what would be the saturation point (after which the performance seizes to improve). Multiple experiments were setup to answer these questions and are detailed below.

To answer the first question, that is, what would happen if the algorithms in the cascade were switched? Firstly, the processing time would increase. This is because now the stronger algorithm would be executed first and it will process all the enrolled users, thus negating one of the major advantages of the system. Secondly, there is no guarantee that the performance of the system will improve as the final stage of the system would be a weak matcher. The following experimental system was utilised to evaluate the performance. The combination of IFE1 and M2 algorithms with WVUid database and FE1 and M1 algorithms with WVUfd database were used as the two stages of the experimental system. The third row of Table 2 shows the results of this system. The results of the iris– fingerprint cascade in terms of FAR, FRR and accuracy are shown in Fig. 12. The results show that not only does the performance of this system drop with respect to the other cascaded classifier-based system it also drops below that of the unimodal system. This is due to the fact that the final stage of the experimental system is a weak matcher and it is unable to provide a correct answer even for the shorter candidate list. This result highlights an important design issue, which is, the selection of algorithms for each stage of the cascade. Improper selection of algorithms for each stage would deteriorate the performance of the system instead of improving it.

To illustrate the improvement achieved in the processing time an average processing time calculation is required. For the sake of proper time calculations two important assumptions are being considered.

Firstly, a real world system is being considered, that is, the enrolment images of both the fingerprints and iris are pre-processed into feature sets before being saved in the enrolment users' database. This means that at runtime when the user is being recognised only the input image is processed through the feature extractor phase and then the features are matched together to generate the matching scores. Secondly, the average processing time for the matching of a single fingerprint image and a single iris image against the enrolled users database is calculated by running the matching for a 1000 times and calculating the average over the total time.

The feature extraction time for the input image is included into the overall average time to properly evaluate the time performance of the proposed system against a strong unimodal biometric system and a standard multimodal biometric system. For fingerprints the average processing time of the ***FE*1** and ***M*1** algorithm combination was evaluated and for iris the average processing time of ***IFE*1** and ***M*2** algorithm combination was evaluated. The times were evaluated on a Windows machine with a dual core 1.7 GHz processor and 2 GB of RAM using MATLAB. The average processing times for matching of a single iris image was evaluated to be about 0.977 s and that for a single fingerprint image was 0.124 s. using these times we can estimate the processing times for unimodal and proposed framework-based systems as follows:

- Time required to match the 100 iris images using $M_2 = 0.977 \times 100 = 97.7$ s.
- Best case (with minimum no of users in the candidate list) time requirement for fingerprint–iris cascade = $(0.124 \times 100) + (0.977 \times 15) = 27.05$ s.
- Worst case (with maximum no of users in the candidate list) time requirement for fingerprint–iris cascade = $(0.124 \times 100) +$

(0.977 × 60) = 71.02 s.

The above calculations clearly show that the proposed framework not only improves the accuracy (while keeping the error rates consistent with the strong unimodal system) it also reduces the processing time by about 27% even in the worst case scenarios.

It is interesting to note that in [7] the authors claim to provide an improved performance over a strong unimodal system (Daugman's) by using a single modality cascaded approach. Sun *et al*. claim to improve the accuracy by about 10%, which is comparable to the improvement provided by our proposed system, but their average time cost increases by 2% over the processing time of the unimodal system, whereas our proposed system reduces the processing time significantly even for the worst case scenarios.

The following experimental system was designed to test the performance of the proposed framework on a single modality. The combinations of ***FE*1** and ***M*1** algorithms and ***FE*2** and ***M*1** algorithms were used for the two stages of the cascaded classifier-based system, respectively. ***WVUfd*** database was used for testing. The results are presented in Table 3. The results show a significant improvement of over the unimodal system.

Finally, it is important to note that the performance of a cascaded classifier-based biometric recognition system will not keep increasing additional classifier stages. A point will be reached when the performance will not increase anymore. This point can be called the saturation point. The addition of more classifier stages beyond the saturation point will result in the increase of error rates and the reduction in accuracy. The saturation point would differ for different systems and will be depended upon the types of modalities being used, the size of the test and implementation databases and the type and number of classification stages being used. For example, for the first experimental system presented in this section, a 100 users based test database is used. In addition, the first stage combination of ***FE*1** and ***M*1** is quite stable; therefore, the system is already at its saturation point, which means that addition of any new classification stage would not show any improvement in performance. To verify this, another classification stage based on the combination of ***FE*2** and ***M*1** algorithms is added to the system and the results of the two stage and three stage systems are presented in Table 4. The results clearly show that even with the addition of the second classification stage of ***FE*2** and ***M*1** the number of genuine matches do not increase but the number of false accepts and false rejects increase. This behaviour indicates the saturation point of the experimental setup. For an optimal performance, the designed system should contain at least one less classification stage then the saturation point. It is important to note that this particular configuration of matcher and datasets used have the saturation point of two stages. Other systems will behave differently and they should be thoroughly tested to evaluate their proper saturation points and operated at number of stages just below that point.

5 Conclusions

This paper discussed a cascaded classifier-based framework for use in biometric recognition system. The proposed framework utilises a set of weak classifiers to reduce the dataset to a small list of candidate users. This list is used in the final stage of the cascade to provide the decision. Weaker classifiers are employed for sections of the cascade responsible for the generation of the candidate lists and a stronger classifier is preferred at the end of the ensemble to calculate the decision. Each classifier in the ensemble can be designed to use a different modality as an input or it can be designed to work with the same modality.

The candidate list is generated based on the quality of the match scores using Mahalanobis distance-based quality measure to compute the quality of matching. The number of side-lobes in the match score quality matrix is used as aparameter to reduce the candidate list. This parameter is able to detect the quality of images in the database as well as identifying the algorithms suitable for use in the cascaded classifier. The experimental results have shown the effectiveness of the proposed framework. The proposed framework also reduces the processing time by reducing the number of enrolled user the strong classifier has to evaluate. This paper finally evaluates the performance of the proposed framework for both single and multiple modalities and explores the saturation point in terms of performance improvement. It was shown that the system performance will not keep increasing as we add more stages. In fact, the system will reach a saturation point in terms of performance very soon and any addition of more stages to the system will only serve to increase the complexity of the whole system.

6 References

1 Schapire, R.: 'Strength of weak learnability', *Mach. Learn.*, 1990, 5, pp. 197–227
2 Freund, Y.: 'Boosting a weak learning algorithm by majority'. Proc. Third Annual Workshop on Computational Learning Theory, 1990
3 Breiman, L.: 'Bagging predictors', *Mach. Learn.*, 1996, 24, (2), pp. 123–140
4 Kearns, M., Valiant, L.G.: 'Learning Boolean formulae or finite automata is as hard as factoring'. Technical Report TR-14-88, Harvard University Aiken Computation Laboratory, August 1988
5 Viola, P., Jones, M.J.: 'Rapid object detection using a boosted cascade of simple features'. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 2001
6 Xu, L., Krzyzak, A., Suen, C.Y.: 'Methods of combining multiple classifiers and their applications to handwriting recognition', *IEEE Trans. Syst. Man Cybern.*, 1992, 22, pp. 418–435
7 Sun, Z., Wang, Y., Tan, T., Cui, J.: 'Improving iris recognition accuracy via cascaded classifiers'. Proc. First Int. Conf. on Biometric Authentication, (*LNCS* 3072), 2004, pp. 418–425
8 Iqbal, A., Namboodiri, A.M.: 'Cascaded filtering for biometric identification using random projections'. Proc. 2011 National Conf. on Communication, 1–5, 2011

9 Sigari, M.H., Pourshahabi, M.R., Pourreza, H.R.: 'An ensemble classifier approach for static signature verification based on multiple-resolution extracted features', Int. J. Signal Process., Image Process. Pattern Recognit., 2012, 5, (1), pp. 21–36
10 Praviz, M., Moin, M.S.: 'Boosting approach for score level fusion in multimodal biometrics based on AUC maximization', J. Inf. Hiding Multimed. Signal Process., 2011, 2, (1), pp. 51–59
11 Erzin, E., Yemez, Y., Takelp, M.A.: 'Multimodal speaker identification using an adaptive classifier cascade based on modality reliability', IEEE Trans. Multimed., 2005, 7, (5), pp. 840–852
12 Lee, P.H., Chu, L.J., Huang, Y.P., Shih, S.W., Chen, C.S., Wang, H.M.: 'Cascading multimodal verification using face, voice and iris information'. IEEE Int. Conf. on Multimedia and Exp., 2007, pp. 847–850
13 Lakshmiprabha, N.S., Bhattacharya, J., Majumder, S.: 'Face recognition using multimodal biometric features'. Int. Conf. on Image Information Processing, 2011
14 Soviany, S., Soviany, C., Jurian, M.: 'A multimodal approach for biometric authentication with multiple classifiers', World Acad. Sci. Eng. Technol., 2011
15 Baig, A., Nawaz, R.: 'Cascaded face recognition system via sparse representation'. Proc. Ninth IEEE Int. Bhurban Conf. on Applied Science and Technology, 2012
16 Baig, A., Bouridane, A., Kurugollu, F.: 'A novel modality independent score-level quality measure'. IEEE, IET Int. Symp. on Communication Systems, Networks and Digital Signal Processing, 2010
17 Mahalanobis, P.C.: 'On the generalised distance in statistics'. Proc. National Institute of Sciences of India, 1936, pp. 49–55
18 Shi, Z., Govindaraju, V.: 'A chaincode based scheme for fingerprint feature extraction', Pattern Recognit. Lett., 2006, 27, pp. 462–468
19 Baig, A., Bouridane, A., Kurugollu, F.: 'A corner strength based fingerprint segmentation algorithm with dynamic thresholding'. Proc. 19th Int. Conf. on Pattern Recognition (ICPR), 2008
20 Niblack, W.: 'An introduction to digital image processing' (Prentice-Hall, Englewood Cliffs, NJ, 1986)
21 Tico, M., Kuosmanen, P.: 'New approach of automated fingerprint matching'. Proc. SPIE Symp. on Electronic Imaging Systems and Image Processing Methods, Real-Time Imaging V, San Jose, California, USA, 20–26 January 2001, vol. 4303, pp. 115–126
22 Daugman, J.G.: 'High confidence visual recognition of persons by a test of statistical independence', IEEE Trans. Pattern Anal. Mach. Intell., 1993, 15, (11), pp. 1148–1161
23 Masek, L., Kovesi, P.: 'MATLAB source code for a biometric identification system based on iris patterns' (The School of Computer Science and Software Engineering, The University of Western Australia, 2003)
24 Chikkerur, S., Cartwright, A.N., Govindaraju, V.: 'K-plet and coupled BFS: a graph based fingerprint representation and matching algorithm'. Int. Conf. on Biometrics, 2006, pp. 309–315
25 Hornak, L.A., Ross, A., Crihalmeanu, S.G., Schuckers, S.A.: 'A protocol for multibiometric data acquisition storage and dissemination'. Problem/Technical Report, West Virginia University, 2007

Figures:

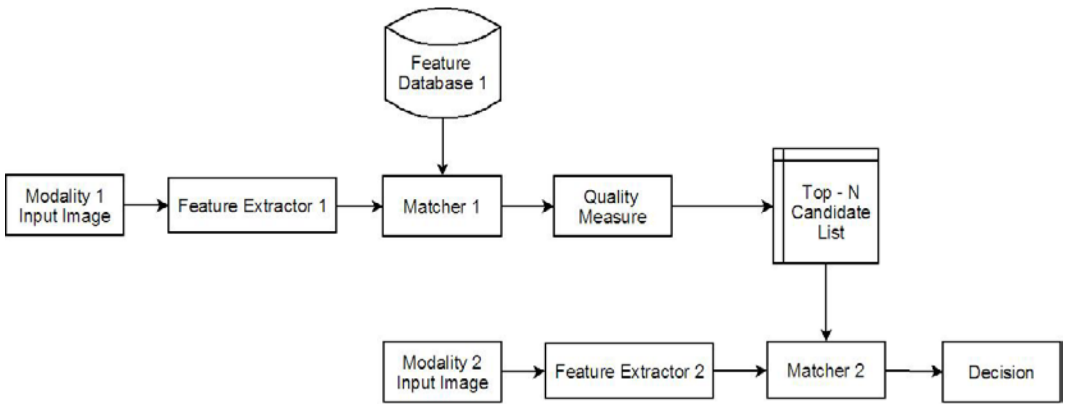

Fig. 1 Proposed cascading multimodal biometric recognition system

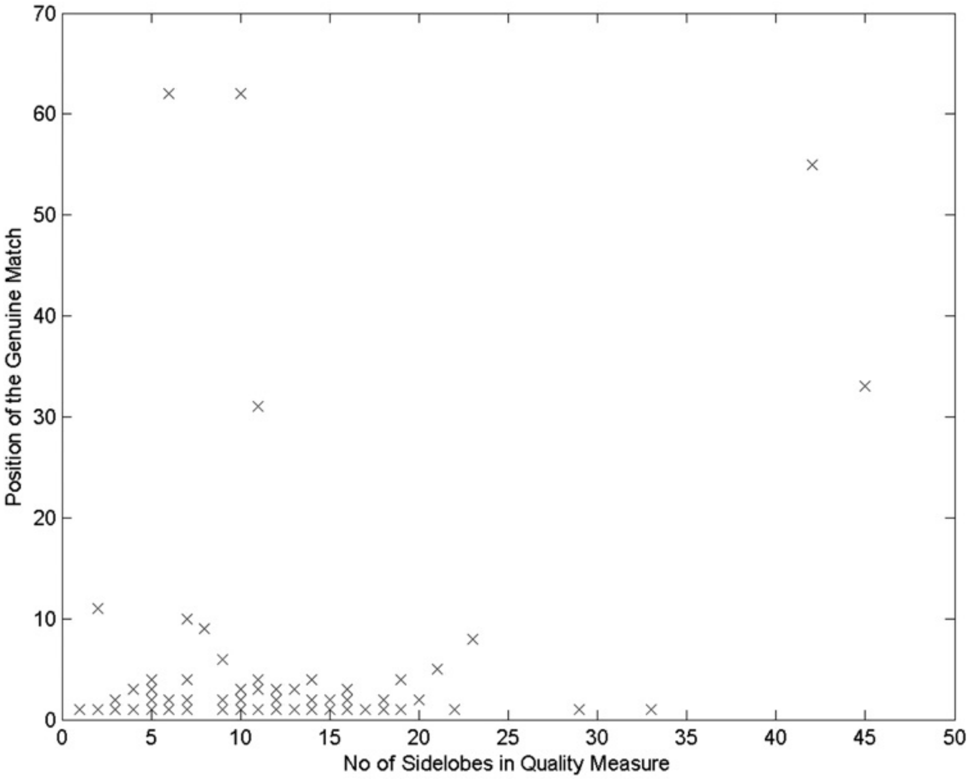

Fig. 2 Plot of number of side-lobes against genuine match position for FE1 and M1

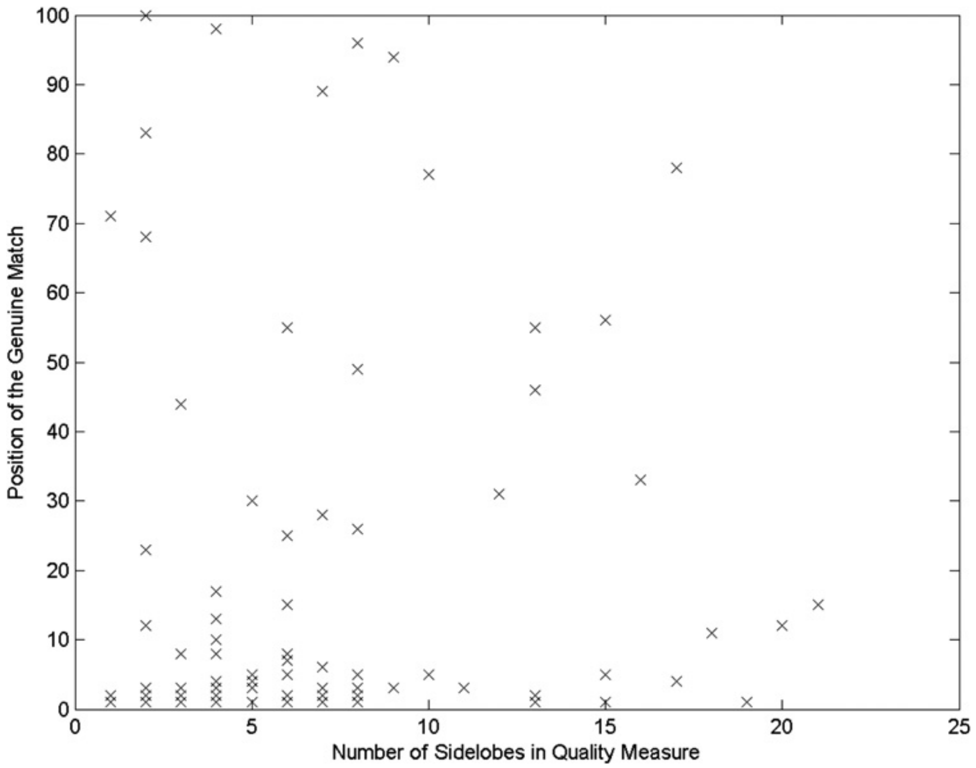

Fig. 3 Plot of number of side-lobes against genuine match position for FE2 and M1

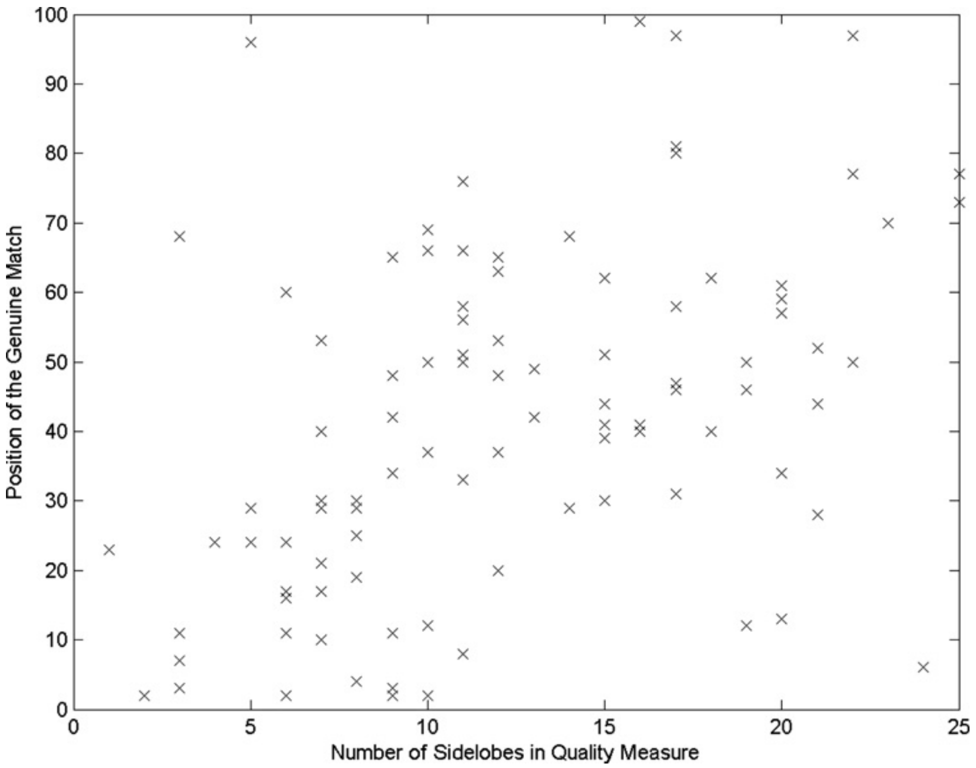

Fig. 4 Plot of number of side-lobes against genuine match position for FE1 and M2


Fig. 5   Example of low-quality input images

Fig. 6 Example of medium quality input image

Fig. 7 Example of good quality input image

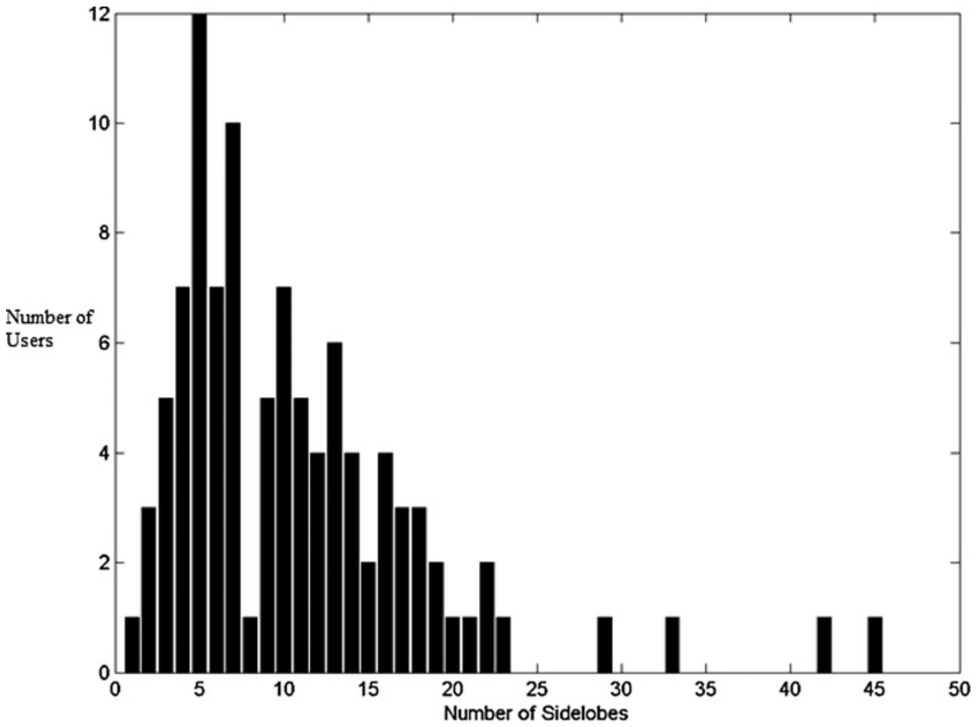

Fig. 8 Plot of number of side-lobes against number of users

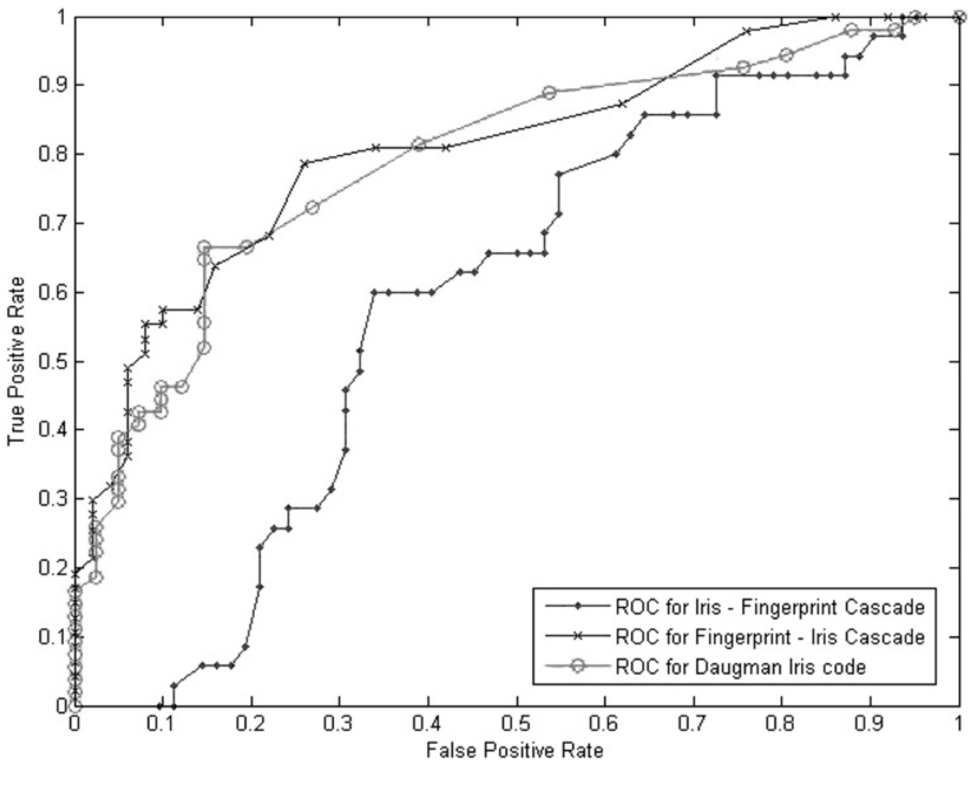

Fig. 9 ROC curve for the three experimental systems

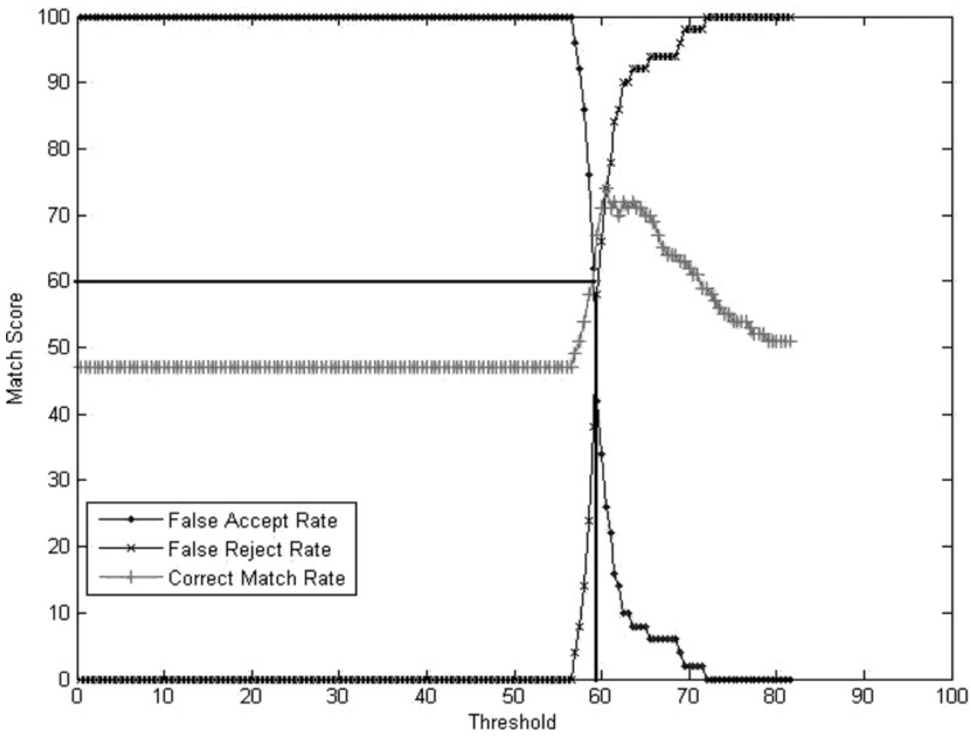

Fig. 10 FAR, FRR and accuracy against threshold for iris recognition system

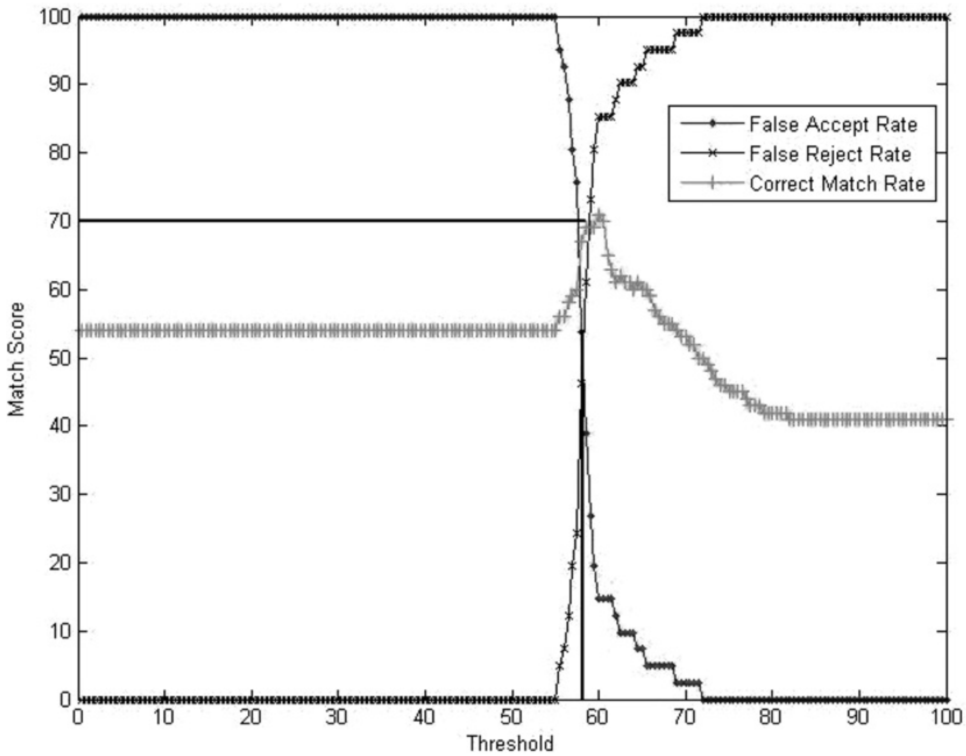

Fig. 11 FAR, FRR and accuracy against threshold for fingerprint–iris cascade

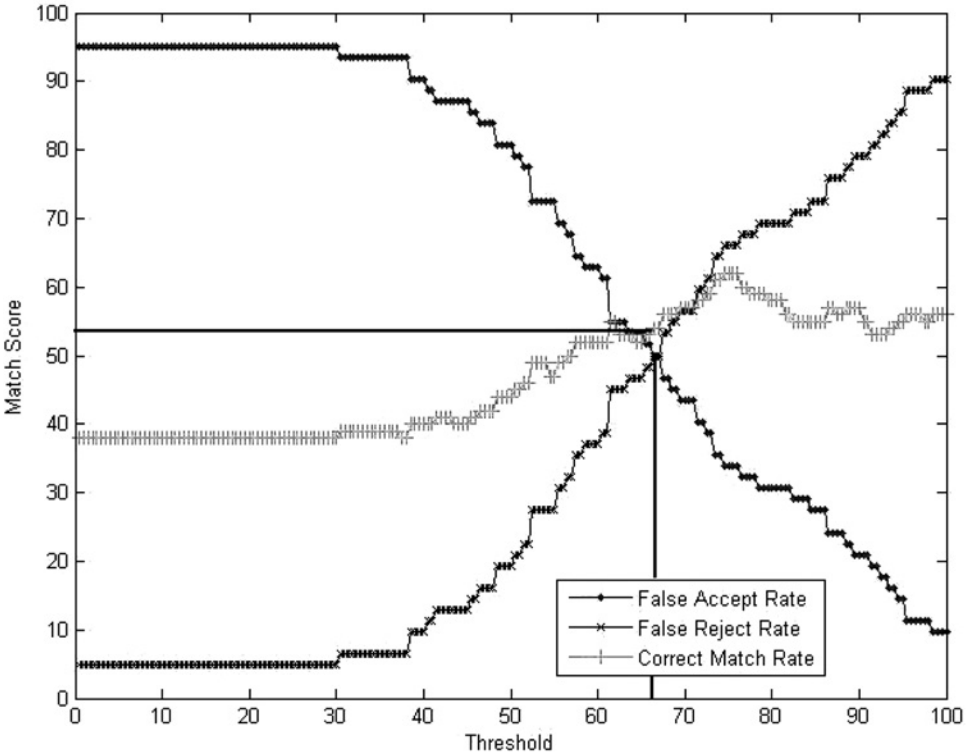

Fig. 12 FAR, FRR and accuracy against threshold for iris–fingerprint cascade

Tables :

Table 1 Candidate list size

Condition based on number of side-lobes	Number of users in candidate list
$0 < \text{side-lobes} \leq 10$	15
$10 < \text{side-lobes} \leq 20$	35
$\text{side-lobes} > 20$	60

Table 2 Experimental results (unimodal against proposed system)

	Genuine match	False accept	False reject	Imposter match
iris system on WVUid	36	12	12	40
proposed system ($FE1 + M1$)+ ($IFE1 + M2$)	40	13	13	31
proposed system ($IFE1 + M2$)+ ($FE1 + M1$)	16	19	19	43

Table 3 Experimental results for single modality

	Genuine match	False accept	False reject	Imposter match
fingerprint system ($FE2 + M1$)	9	21	21	49
proposed system ($FE1 + M1$)+ ($FE2 + M1$)	26	17	17	40

Table 4 Experimental results for three stage system

	Genuine match	False accept	False reject	Imposter match
iris system on WVUid WVUid	36	12	12	40
proposed system ($FE1 + M1$)+ ($IFE1 + M2$)	40	13	13	31
three stage proposed system ($FE1 + M2$) + ($FE2 + M1$) + ($IFE1 + M2$)	40	14	14	27