Multibiometric Secure Index Value Code Generation for Authentication and Retrieval

K.Saranya[1] K.Baskar[2]
[1]M.E [2]Assistant professor
[1, 2] PGP College of Engineering and Technology, Namakkal-637207

Abstract— The use of multiple biometric sources for human recognition, referred to as multibiometrics, mitigates some of the limitations of unimodal biometric systems by increasing recognition accuracy, improving population coverage, imparting fault-tolerance, and enhancing security.
In a biometric identification system, the identity corresponding to the input data (probe) is typically determined by comparing it against the templates of all identities in a database (gallery).
An alternative e approach is to limit the number of identities against which matching is performed based on criteria that are fast to evaluate. We propose a method for generating fixed-length codes for indexing biometric databases. An index code is constructed by computing match scores between a biometric image and a fixed set of reference images. Candidate identities are retrieved based on the similarity between the index code of the probe image and those of the identities in the database. The number of multibiometric systems deployed on a national scale is increasing and the sizes of the underlying databases are growing.
These databases are used extensively, thereby requiring efficient ways for searching and retrieving relevant identities. Searching a biometric database for an identity is usually done by comparing the probe image against every enrolled identity in the database and generating a ranked list of candidate identities. Depending on the nature of the matching algorithm, the matching speed in some systems can be slow.
The proposed technique can be easily extended to retrieve pertinent identities from multimodal databases. Experiments on a chimeric face and fingerprint bimodal database resulted in an
84% average reduction in the search space at a hit rate of 100%. These results suggest that the proposed indexing scheme has the potential to substantially reduce the response time without compromising the accuracy of identification. New representation schemes that allow for faster search and, therefore, shorter response time are needed.
Keywords: Biometrics, feature extraction, image retrieval, indexing, pattern matching.

I. INTRODUCTION

THE use of multiple biometric sources for human recognition, referred to as multibiometrics, mitigates some of the limitations of unimodal biometric systems by increasing recognition accuracy, improving population coverage, imparting fault-tolerance, and enhancing security. The number of multibiometric systems deployed on a national scale is increasing and the sizes of the underlying databases are growing. These databases are used extensively, thereby requiring efficient ways for searching and retrieving relevant identities. Searching a biometric database for an identity is usually done by comparing the probe image against every enrolled identity in the database and generating a ranked list of candidate identities. Depending on the nature of the matching algorithm, the matching speed in some systems can be slow. New representation schemes that allow for faster search and, therefore, shorter response time are needed.
The retrieval of a small number of candidate identities from a database based on the probe data is known as database filtering. Filtering can be accomplished by using classification or indexing schemes. In a classification scheme, identities in the database are partitioned into several classes. Only the identities belonging to the same class as that of the probe image are retrieved during the search process for further comparison. This approach has two main Limitations:

1) It assumes that each identity can be unambiguously assigned to a single class; and 2) the distribution of identities across classes may be uneven resulting in inefficient classification. We present a method for indexing multimodal biometric databases based on index codes generated by a biometric matcher. The indexing mechanism is executed separately for each modality and the results are combined into a final list of potential candidates the proposed indexing technique relies on the use of a small set of reference images for each modality.
A modality-specific index code is generated by matching an input image against these reference images, resulting in a set of match scores. During identification, the index code of the input image is compared to the index codes of the enrolled identities in order to find a set of potential matches. The index codes of multiple modalities are fused to improve the accuracy of indexing resulting in a robust and efficient indexing system. This approach relies on a matcher, which is an integral part of every automated biometric identification system. Because the generated index codes are compact and their (dis)similarity can be computed rapidly, the approach has low storage requirements and can improve the system response time even for small databases.
This manuscript is organized as follows.
Section II presents a brief review of previous research on indexing and classification of biometric databases. Identification techniques that use reference data are also discussed. Section III describes the proposed indexing methodology and how it improves computational time. Indexing multimodal databases using two different

fusion techniques is discussed in Section IV. The effect of various indexing parameters on overall performance is studied in Section V. The experimental configuration is summarized in Section VI. The performance of the proposed scheme is resented in Section VII. Section VIII includes conclusions and directions for future work.

II. RELATED WORK

A. *Fingerprint Indexing*

The problem of fingerprint classification has been studied extensively. The first published fingerprint classification method, the Henry classification system [15], was based on the spatial configuration of singular points present in fingerprint patterns. The geometric properties of triangles constructed from minutiae points were first utilized to index fingerprints. Then Fingerprint indexing using ridge orientation was proposed by Lumini et al. [19] and Cappelli et al. [20]. More recent techniques exploit both minutiae points and ridge orientation for indexing fingerprints [21], [22].

The indexing methods listed above require the application of image processing techniques which are specific to the indexing method. To avoid the complexity of designing new feature extraction routines. They adopted a sequential search process in which filtering was performed based on the correlation between the set of match scores that were already computed for the probe and the corresponding match scores for the images in the database that were not yet visited. In this technique, a matrix that contains the pair wise match scores of all images in the database has to be permanently stored and updated for each newly enrolled identity. A drawback of this approach is that storing the matrix of match scores for a database containing millions of images can be impractical.

B. *Face Indexing*

Indexing methods for face databases usually focus on a specific recognition algorithm. The approach is based on the classical eigenface method and uses the coefficients of projection to rank the database images with respect to each eigenface. The probe is ranked in the same way and a local search is performed for each eigenface to find the database image that is closest to the probe. Thus, the reduction in the search space depends on the number of eigenfaces used.

Another approach for face indexing relies on the use of small binary images obtained from an edge detection filter. The goal is to reduce dimensionality and allow faster comparisons during the search-and- retrieval stage. This technique was first proposed who applied the Sobel operator on mug-shot face images and used a modified Hausdorff distance to compare the resulting binary images. Compact binary codes were also used for fast retrieval of face images from online databases.

A combination of classification and indexing was proposed by Perronnin and Dugelay [27]. A face database was split into a predefined number of classes by applying a clustering technique on parametric models of the enrolled faces. The low- dimensionality of the index vectors and the fast computation of the similarity metric reduced the computational cost for the overall identification process but with a 5% loss in identification accuracy.

C. *Indexing Using Match Scores*

Reducing the dimensionality of the biometric template by using a set of distance (or similarity) scores to a fixed set of templates has been used in speaker recognition. Sturim et al. [28] Compared each enrolled speaker against a fixed set of speaker models (called anchor models). The resulting set of scores was used as a projection of the speaker in the space defined by the anchor models. Retrieval was based on the distances between the probe and the enrolled speakers in the projection space. The calculation of these distances is much faster compared to conventional speaker matching schemes that utilize Gaussian mixture models.

We propose an indexing method that is similar to the anchor models approach and can be applied to any biometric modality. This is achieved by using the matcher inherent in the biometric system to create index codes (index vectors). Dimensionality reduction is not applied to the index codes. Instead, a small number of reference images, which form the basis of the index codes, are selected to ensure that the variance among the index codes is large.

In contrast to the original approach of anchor models, our method uses raw image data and does not require training. Besides being applicable to any biometric modality, this indexing method can be easily incorporated into any existing biometric identification system. We demonstrate how the information available in multimodal biometric databases can be used to achieve fast retrieval and low error rates, even when each individual is enrolled with a single image for each modality. The creation of an index code involves matching the input image to a set of reference images.

III. INDEX CODES FROM IMPOSTOR MATCH SCORES

The proposed indexing technique can either employ the biometric matcher that is already present in the biometric system or use another independent matcher. Index codes are generated for each modality using the corresponding matcher.

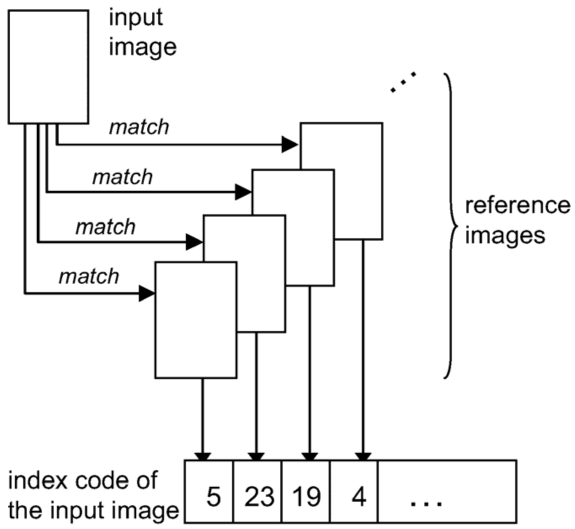

Fig. 1: Generation of an index code

During retrieval, the index code of the probe is compared against those in the gallery using a similarity measure to retrieve a list of candidate identities for biometric matching.

A. *Indexing a Single Modality*

In this section, the face modality is used as an example to illustrate the process. However, the inferred properties are applicable to the fingerprint modality (as observed in our experiments) and perhaps to other biometric modalities as well.

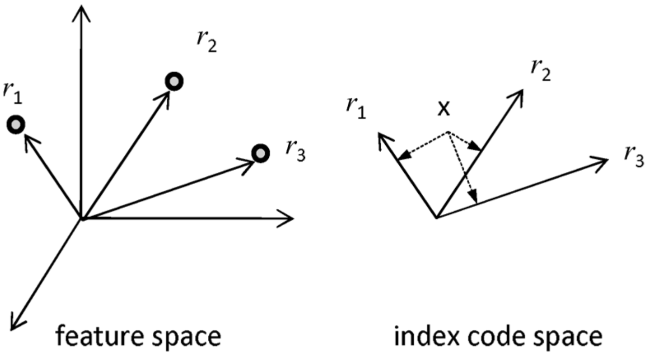

Fig. 2: Geometric interpretation of the proposed indexing approach

The reference images may be viewed as "basis" vectors in the original feature space.
The number of reference images for each modality can be different. The index codes for each enrolled identity are stored in the database and used in a fusion framework during the retrieval process.

B. *Conditions to Achieve Speedup*

The retrieval process performs an exhaustive search across the index codes of the enrolled identities. Thus, an improvement in the speed of identification is possible only if the search space is substantially reduced and if the distance between two index codes can be computed in a fraction of the time needed to match two biometric templates.
Let P be the fractional reduction in the number of candidate identities achieved by the indexing scheme when applied on a database of size M. Let denote the dimensionality of the index code.
The overall computation time of the identification system can be approximated by the sum of the matching operations between the input image and the reference images, the M operations for computing the distances between the index codes of the probe and the enrolled identities, and the P*M matching operations required for the final identification. Similarly, the time needed for identification without indexing consists of M matching operations.

IV. INDEX CODES FOR MULTIMODAL DATABASES

There is an inherent trade-off between the total number of retrieved candidates and the number of correctly retrieved candidates. Fusion schemes are often useful for narrowing down the total number of retrieved candidates and/or increasing the number of correctly retrieved candidates. Index codes are stored separately for each modality thereby making the indexing scheme flexible in including more modalities or excluding a certain modality. The ability to exclude a modality from the indexing process is valuable when prior knowledge indicates that a certain modality is unreliable or when data for a modality are missing. Our general approach for indexing multimodal databases is shown in Fig.3

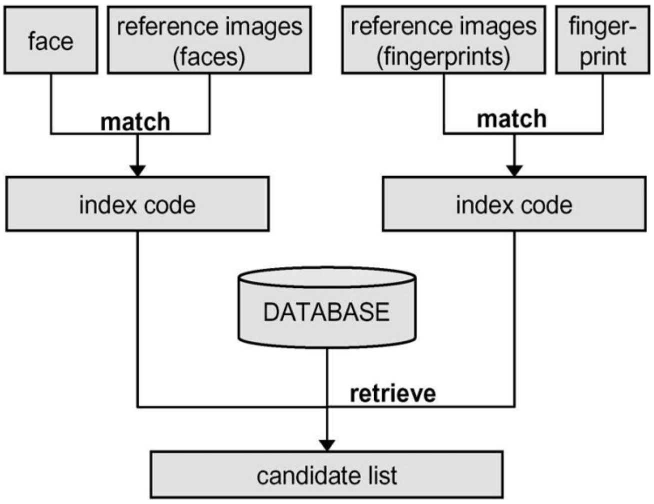

Fig. 3: Indexing two modalities

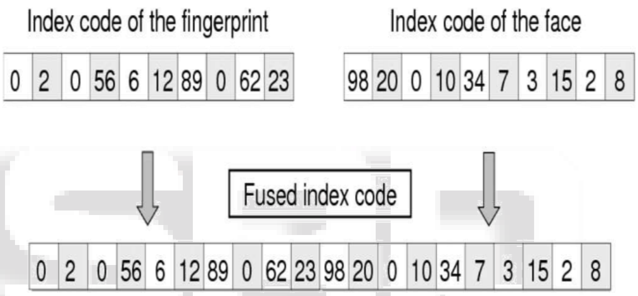

Fig. 4: Fusion by concatenation of index codes

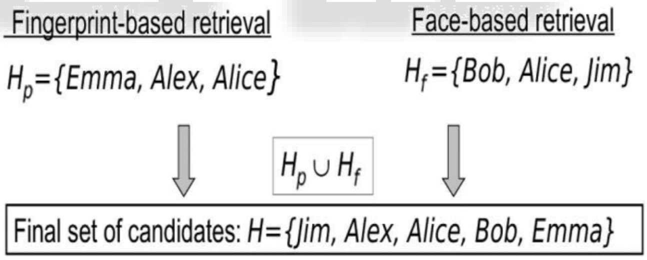

Fig. 5: Fusion by union of candidate lists.

This fusion scheme results in longer index codes. Ideally, using longer index codes results in larger variances among them—this is desirable. One weakness of this fusion scheme is that poor indexing performance due to one of the modalities can negatively affect the overall performance of indexing.

V. PARAMETERS OF THE PROPOSED INDEXING SCHEME

A. *Similarity Measures for Index Codes*

Although most data collection protocols impose strict constraints on the data acquisition process, noise in the input images can significantly impact the match scores and, consequently, the index codes. The association between two index codes can be measured by their correlation. Index codes belonging to the same identity are expected to have a strong positive correlation. Index codes belonging to different identities are expected to be uncorrelated.

B. Dimensionality of the Index Codes

While using a larger number of reference images can improve indexing performance it also increases the computational requirements of the method (as discussed in Section III-B). Furthermore, increasing the number of reference images beyond a certain number is not beneficial because the improvement in accuracy will be insignificant compared to the increased overhead. Generally, as more images are included in the reference set, the variability among them decreases (unless the biometric template has infinite capacity). Therefore, this number should be chosen empirically according to the desired accuracy and speedup. We provide guidelines for choosing this number in Section VII.

C. Selecting Reference Images

Reference images can be selected from the database itself. They can also be synthetically generated images. While the entire database can be viewed as a candidate pool for selecting reference images, practical considerations dictate the use of a small random subset of images for this purpose. A greater degree of diversity among the reference images increases the probability that the index codes of different subjects will be unique and well-spread in space. We consider three different selection rules for ensuring good diversity.

First, the max-variation rule selects reference images with the largest variances of impostor match scores.

Second, the max-mean rule selects images whose impostor match scores have a large mean value.

Third, the min-correlation rule selects an optimal set of reference images by 1) starting with the entire candidate pool, 2) removing the image whose average correlation to other images in the set is the highest, and 3) repeating this process until the desired number of reference images is obtained.

Thus, some of the selected reference image may have very similar characteristics. While this phenomenon does not necessarily reduce the hit rate, it results in redundant entries within each index code. The min-correlation rule attempts to overcome this drawback by reducing the pair wise correlation among the impostor match scores of the reference images

VI. EXPERIMENTS

There are very few publicly available multimodal biometric databases. Examples include WVU [34], Bio Secure [35], XM2VTS [36], MBGC [37], and BANCA [38]. However, these databases have small numbers of subjects therefore, cannot be used to evaluate our indexing approach in a reliable manner. Therefore, we assembled a chimeric multimodal dataset using the FERET face database [39] and the WVU fingerprint database [34]. There are 1195 subjects with frontal face images in the FERET database. We used only 1010 of these subjects because the images of the remaining 185 subjects could not be processed by the face matcher used in this work. Sample images are shown in Fig.6.

The WVU fingerprint database contains images of 4 different fingers (left index, left thumb, right index, right thumb) from 270 subjects. We treated the individual fingers as independent "subjects," resulting in a total of 1080 subjects. However, because the matcher could not process the images of 210 subjects, a total of 870 subjects were used in the experiments.

Fig. 6: Sample images from (a) the FERET and (b) the FRGC databases. Faces with smiling expressions were enrolled in the database, while those with neutral expressions were used as probes for evaluating performance

Fig. 7: Sample images from the WVU fingerprint database. Images from different fingers of the same individual were treated as different subjects

The WVU fingerprint database contains images of 4 different fingers from 270 subjects. We treated the individual fingers as independent "subjects," resulting in a total of 1080 subjects. However, because the matcher could not process the images of 210 subjects, a total of 870 subjects were used in the experiments. Two images per subject were used from the WVU database—one for enrolling the subject into the database and the other one as a probe image. Sample images from the WVU database are shown in Fig. 7.

VII. RESULTS

Unless otherwise specified, the results in this section were obtained by performing 10-fold cross-validation repeated 10 times, which gives 100 estimates of the penetration rate for a given hit rate. The mean value and the 99th percentile from the distribution of these estimates were used to assess the performance of our indexing approach.

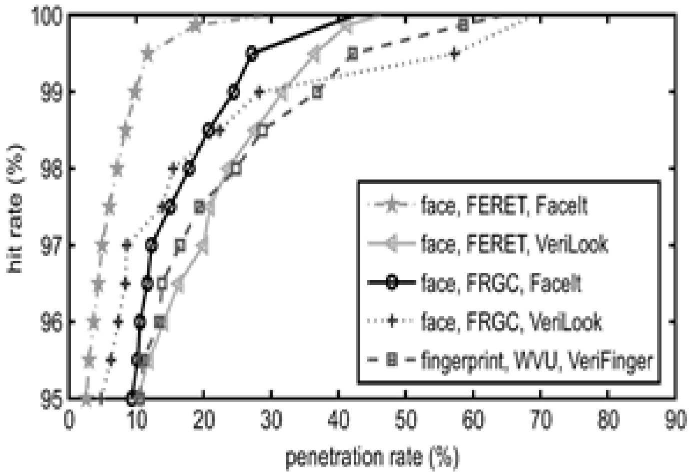

Fig.8: Indexing performance of single modalities using different databases and matchers.

Results for each database were obtained by using reference images from the same database. For the face modality, the reference set was determined by using the scores produced by the Veri Look and the same set was used when evaluating the performance using the Face It matcher

A. Unimodal Databases

The performance of the proposed indexing method on the two faces databases and the fingerprint database. Performance was better for the face modality than for the fingerprint modality. The main reason for this result is probably the large number of zero-valued match scores in the fingerprint modality. The set of fingerprint scores generated by matching an image in the WVU database against other images in the database contains 27% zero-valued scores, on average.

B. Choice of Matcher

The choice of face matcher had a strong effect on the performance of indexing. The FaceIt (FI) matcher resulted in consistently lower penetration rates compared to the VeriLook (VL) matcher, which can be seen in Fig. 12. The FaceIt matcher also has better recognition performance, which might be the reason for its better indexing performance.

C. Reference Images From a Different Database

Changes in illumination conditions and pose of the head are common problems in face recognition. Similarly fingerprint recognition has larger error rates when the enrolled image and the probe are captured by different sensors. Thus, it is logical to select the reference images from the database of enrolled images.

D. Bimodal Databases

Poh and Bengio [41] stated that, in certain cases, identification results obtained on chimeric multimodal databases may not be representative of the true identification performance. To account for these situations, the performance of multimodal indexing was evaluated by extensive cross-validation. Thus, the chance of underestimating the true (unknown) penetration rate was reduced. In this experiment, the face and fingerprint images from the FERET and WVU databases, respectively, were paired in an exhaustive manner.

E. Effect of the Size of the Database

A simulation study was conducted to assess the performance of the proposed system on large databases (i.e., containing thousands of identities). Using the FERET database, a large number of index codes of dimension were created by modeling the impostor match scores of the face matcher. The new index codes were generated by randomly sampling the set of match scores associated with the reference images.

VIII. CONCLUSION & FUTURE WORK

The proposed technique is not modality-specific. Therefore, it can easily be incorporated into existing biometric systems. The biometric matcher that is inherent to the system can be used for generating index codes. Furthermore, the application of our approach to multi biometric databases is straightforward. Using the proposed indexing technique on a chimeric multimodal database resulted in a reduction of the search space by an average of 84% at a 100% hit rate. The use of reference images that had different sizes, image resolutions, and color depths, compared to the images in the database, did not change the performance of the proposed indexing method substantially. In this case, penetration rates were higher only for hit rates above 99%. Results from indexing a chimeric bimodal database indicated that fusion by union of candidate lists had better performance than fusion by concatenation of index codes. Z-score normalization played an important role in optimizing the performance of the two fusion techniques. The main factor for the amount of speedup during identification was the penetration rate of the indexing.

IX. REFERENCES

[1] A. Gyaourova and A. Ross, “A coding scheme for Indexing multimodal biometric databases”, Jun. 2009.

[2] T. Vikram, C. Gowda, D. Guru, and S. Urs, “Face indexing and retrieval by spatial similarity,” May 2008.

[3] R. J. Bayardo, Y. Ma, and R. Srikant, “Scaling up all pairs similarity search,” in Proc. 16th Int. Conf. World Wide Web, New York, 2007.

[4] A. Ross, K. Nandakumar, and A. Jain, ”Handbook of Multibiometrics”, 1st ed. New York: Springer, 2006.

[5] A. Mhatre, S. Chikkerur, and V. Govindaraju, “Indexing biometric databases using pyramid technique,”, Jul. 2005.

[6] A. Mhatre, S. Palla, S. Chikkerur, and V. Govindaraju, “Efficient search and retrieval in biometric databases,” 2005.

