Multi-biometric Templates Using Fingerprint and Voice

Eren Camlikaya, Alisher Kholmatov, Berrin Yanikoglu

Faculty of Engineering and Natural Sciences
Sabanci University, Istanbul, 34956, Turkey

ABSTRACT

As biometrics gains popularity, there is an increasing concern about privacy and misuse of biometric data held in central repositories. Furthermore, biometric verification systems face challenges arising from noise and intra-class variations. To tackle both problems, a multimodal biometric verification system combining fingerprint and voice modalities is proposed. The system combines the two modalities at the template level, using multi-biometric templates. The fusion of fingerprint and voice data successfully diminishes privacy concerns by hiding the minutiae points from the fingerprint, among the artificial points generated by the features obtained from the spoken utterance of the speaker. Equal error rates are observed to be under 2% for the system where 600 utterances from 30 people have been processed and fused with a database of 400 fingerprints from 200 individuals. Accuracy is increased compared to the previous results for voice verification over the same speaker database.

Keywords: multimodal biometrics, multi-biometrics, voice, fingerprint, privacy

1. INTRODUCTION

Due to increasing cyber-memberships and security threats, personal identification and verification have gained significance over the last years. Identification or verification of a claimed identity can be based on 3 major themes; "what you have", "what you know" or "who you are". Themes of "what you know" and "what you have" are quite popular in most of the security scenarios. A simple example of a credit card and its password can be given for such a fusion of the two themes. Systems that are based on "who you are" can be classified as biometric systems which commonly utilize iris, voice, face or fingerprint recognition. These systems can depend on a single behavioral or physiological trait, as well as a combination of them.

Combinations of biometric traits are mainly preferred due to their lower error rates. Using multiple biometric modalities has been shown to decrease error rates, by providing additional useful information to the classifier. Fusion of these behavioral or physiological traits can occur in various levels. Different features can be used by a single system or separate systems can operate independently and their decisions may be combined.[23]

Physiological traits such as fingerprint, iris or hand geometry do not change significantly in time, and more importantly, cannot be changed at will, in contrast to behavioral traits such as voice, signature or gait. Since physiological traits are stable and are unique across individuals, they can be used to track people if biometric databases are misused. Another related concern is that once compromised, a physiological biometric (e.g. fingerprint) cannot be canceled. The privacy and cancelability concerns lead researches to find new solutions, often combining cryptography and biometrics, in recent years.[13, 19, 27]

In this paper, fingerprint and voice data are fused at the template level. We have used fingerprint and voice modalities as two of the most practical and commonly accepted biometrics; in particular, voice is a familiar way of communicating for humans. The gain obtained through the combination is three-fold: increase in privacy, increase in security and cancelability. The constructed multi-biometric templates are *non-unique*,[27] reducing privacy concerns. The performance of the biometric system is increased since multiple biometrics are used in verification. Finally, changing the password in text-dependent voice verification systems can provide cancelable

Further author information: (Send correspondence to Berrin Yanikoglu)
Eren Camlikaya: e-mail: camlikaya@su.sabanciuniv.edu
Alisher Kholmatov: e-mail: alisher@su.sabanciuniv.edu
Berrin Yanikoglu: e-mail: berrin@sabanciuniv.edu, web: http://people.sabanciuniv.edu/berrin

Biometric Technology for Human Identification V, edited by B. V. K. Vijaya Kumar, Salil Prabhakar, Arun A. Ross
Proc. of SPIE Vol. 6944, 69440I, (2008) · 0277-786X/08/$18 · doi: 10.1117/12.777738

biometric templates. Notice that this is not possible with physiological biometrics such as fingerprints. Thus, combining voice, or in general a cancelable biometric, with non-cancelable biometrics is advantageous on several counts.

2. PREVIOUS WORK

Fingerprints have long been used for person verification and identification due to their immutability and uniqueness. Minutiae-based verification approaches are the most common, compared to ridge-based and correlation-based techniques.[17] The performance of minutiae-based fingerprint verification systems heavily depend on the minutiae extraction process done before minutiae alignment. Minutiae extraction is done using image processing operations that take advantage of the rich information available in the ridge structure of a fingerprint. Minutiae alignment, on the other hand, has to be done efficiently and should handle the non-linear deformations present in fingerprints. Jiang and Yau use local structure to find the correspondence between two minutiae sets. They tested their method on a private database of 188 users and achieved an EER under 1%.[12] Jain et al proposed an alignment based algorithm were the ridge information is employed to align the minutiae sets and a bounding box was proposed to match aligned minutiae.[10] Further improvements for this method have been proposed by He et al where the EER is decreased from around 3-4% to 1-2% in a database of 100 users.[8] Tico and Kuosmanen employed orientation field information of the fingerprint pattern to create a fingerprint representation scheme[25] and Ratha et al proposed a matching technique based on the graph representations of the query and template fingerprints, constructed using their respective minutiae features.[19]

Voice is a behavioral biometric which can be used in identity verification, especially over-the-phone applications such as banking. In literature, various text-independent methods have been proposed, some of which are quite successful with equal error rates under 2% using private databases.[7,22] On the other hand, text-dependent systems provide the flexibility of changing the biometric by changing the spoken text. Bellegarda et al introduced a text-dependent system by using singular value decomposition for spectral content matching and dynamic time warping for temporal alignment of the utterances. They have achieved an EER of 4% in a database of 93 speakers.[2] Li et al proposed a method employing Hidden Markov Models for the alignment of utterances with a global phoneme model and achieved an EER of 2.6% with a database of 100 speakers.[15] The feature extraction method used in this paper is based in part on the work by Monrose et al which generates cryptographic keys from voice.[18]

High security applications require very low error rates and unimodal biometric systems are not always satisfying in that regard. In those cases, multi-modal biometric systems are useful. In literature, the combination of multiple biometrics mostly take place at the matching score or decision level.[17,24] Brunelli and Falavigna used the hyperbolic tangent for normalization and weighted geometric average for fusion of voice and face biometrics.[4] These modalities have also been fused by Ben-Yacoub et al by considering several strategies such as support vector machines, tree classifiers and multilayer perceptrons.[3] Kittler et al have experimented with fusion techniques for face and voice on the matching score level.[14] Hong and Jain proposed an identification system using face and fingerprint where the database is pruned via face matching before fingerprint matching.[9]

Privacy is another main concern in building biometric systems, besides low error rates. Numerous architectures have been proposed in recent years, aiming to protect biometric templates stroed in central repositories.[6,13,16,20] Among those, *fuzzy vault* technique is one of the most widely used method where the fingerprint minutiae points are stored with randomly generated chaff points.[13] A user has to provide a certain number of minutiae points to unlock the vault created by the reference fingerprints minutiae set during the enrollment session. Yanikoglu and Kholmatov proposed another method based on the idea of combining multiple biometrics in order to increase both privacy and security.[27] Specifically, minutiae points from two distinct fingers of the same person were combined to create a *multi-biometric template* which is later shown to be more unique, hence more privacy preserving. They also showed that the system provides higher level of security as well, as it verifies both fingerprints. In this paper, we present an implementation of the multi-biometric template framework with fingerprint and voice modalities.

3. PROPOSED METHOD

In this paper we implement the multi-biometric template framework using fingerprint and voice-modalities. In contrast to the original implementation using two fingerprints,[27] this implementation also provides cancelability of the combined template. Voice and fingerprint data of individuals are fused at the template level by combining minutiae points and artificially constructed points obtained from the utterance, as described in the following subsections.

3.1. Feature Extraction from Fingerprint

Minutiae points from the ridge endings and bifurcations on the fingerprint pattern are used as features in our work. In literature, there are several methods proposed for automatic minutiae extraction,[11, 21] which commonly follow well-known image enhancement, binarization and thinning steps. Automatic detection of minutiae points can sometimes result in missed or spurious minutiae. In some systems, minutiae points found through image processing operations are later verified using various post-processing techniques[26] in some systems. After minutiae extraction, finger print verification involves minutiae alignment and matching. In that process, the challenges are caused by non-linear deformations of the fingerprint, as well as missing and spurious minutiae.

Since the aim of this work is to build a multimodal biometric system with information fusion at the template level, we preferred to utilize fingerprint images with manually labeled minutiae points. This is done in order to reduce errors which might arise from the minutiae extraction process. For template generation, the minutiae points from users are stored in a 2 dimensional plane with their x and y coordinates as features.

3.2. Feature Extraction from Voice

The features employed in speaker recognition systems should successfully be able to define the vocal characteristics of the speaker and distinguish it from the voices of other speakers. Short spectra of speech signals give information about both the spoken words and the voice of the speaker. For this purpose, mel frequency cepstral coefficients (MFCCs) utilize the logarithmic energies of the speech data after being filtered by nonuniform frequency filters. The reason for using these filters is that they correspond to the human hearing system the best. Then, discrete cosine transform is applied to the filtered speech data for further decorrelation of the features.[28]

After all the spoken passwords are collected from the speakers, each utterance of every speaker is divided into 30ms frames with 10ms overlap and cepstral analysis is applied to each frame. This means, each 30ms frame is represented by a 12 dimensional vector $< c_1, ..., c_{12} >$ consisting of MFCCs.

Since speech signals can vary in length, each utterance is then aligned with a Hidden Markov Model (HMM) of the corresponding password, in order to determine the correspondence between individual frames and phonemes. The HMM used for this alignment is obtained by concatenating previously trained, speaker- and text-independent phoneme models corresponding to the phonemes of the password. This way, each frame in each utterance is identified as one of the phonemes that may occur in the utterance of digits ([0-9] and "oh"), or noise or silence. The global phonetic Hidden Markov Models used for the alignment are 3-stage, monophone phonetic models which were previously trained using voice samples collected from various users. Phonetic HMMs are commonly used in automatic speech recognition systems, as versatile alternatives to word HMMs.

After this alignment, frames which correspond only to the middle (2nd) stage are kept while the first and final (1st and 3rd) stages of phonemes are deleted. This step is done to reduce the effects of noise and speech variations. The 3-stage phonetic models described above and the alignment process are illustrated in Figure 1.

At this point, mean vectors of cepstral coefficients for each phoneme are calculated by averaging the 12 dimensional vectors representing the frames within the same phoneme (middle stage). Hence, the n^{th} segment (middle stage of a phoneme) is represented by a 12-dimensional mean vector F_n. During the training and testing phases for the system, mean vectors of the phonemes will be used instead of single frame vectors. In the database used, there were on average 3 frames in the middle stages of the phonemes and the total number of segments in an utterance is around 25 for a 6-digit password.

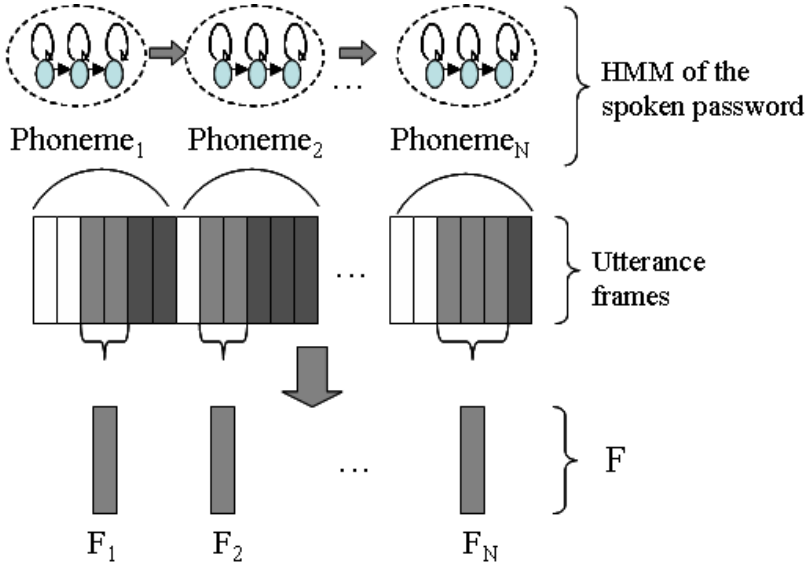

Figure 1. Alignment with 3-stage HMMs: Previously trained 3-stage phonetic HMMs are used in aligning an utterance, to find the correspondence between individual frames and phonemes. Phoneme 1-N indicate the phonemes that occur in spoken password. Levels of grey (white-grey-dark grey) indicate the 3-stages within a phoneme.

3.2.1. Multi-biometric Template Generation

After the feature extraction process, 12-dimensional mean vectors of each segment are concatenated to form a vector F to represent the entire utterance. Therefore, the F vector of an utterance with N segments is $N \times 12$ dimensional. Next, the F vector representing an utterance is binarized using a global threshold t of -3.5 for females and -1 for males. The binarization process is done by assigning bits to each dimension of the F vector according to its value (if the MFC coefficient is above the threshold, 1 is assigned to that dimension). This way, every utterance is now represented with a bit string B with length $N \times 12$. The t values are determined in order to have equal number of 1 and 0 bits for all speakers. Binarization process is illustrated in the top two rows of Figure 2.

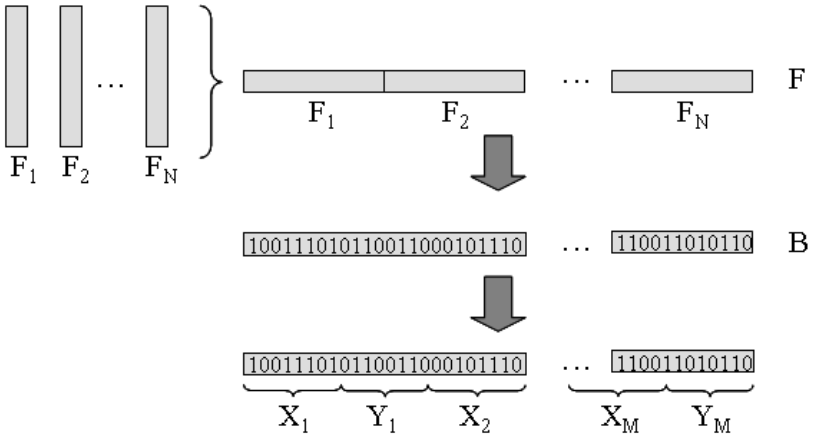

Figure 2. Minutiae point generation from voice: mean feature vectors from the previously aligned utterances are concatenated and binarized according to a predetermined threshold, then the bit string is divided into chunks of 8 bits to obtain the utterance points (X_i, Y_i).

For using voice as a biometric, it is necessary to go through a speaker specific enrollment session. In this paper, there is only one enrollment session where 10 utterances of the chosen password are collected from each user, whereas there may be multi enrollment sessions distributed over time to form a more general spoken password templates.[22] Among the collected utterances, 8 out of 10 utterances from speakers are used for enrollment and the remaining 2 utterances are used for testing. The binary feature vectors B_i, generated from each of the 8 enrollment utterances of a user, are combined so as to obtain a single feature vector, by majority voting of the bits in every dimension. This single feature vector represents the voice template of a single speaker, dependent on the chosen password.

Finally, to map the resulting binary voice template onto the 2-dimensional space of minutiae points, we followed a few different methods with roughly similar results: in one, we divided each binary feature descriptor of a users into groups of 16 and used each 16 bits to generate one point (X,Y) in the 2-dimensional space. This process is shown in the bottom row of Figure 2. These points will be called the "utterance minutiae" of users from now on. To complete the enrollment phase, minutiae points extracted from the fingerprint (A) are combined with the utterance minutiae (B) of the user to form the multi-biometric template (A+B) for the user, *without* any labels indicating the origin of the points. Fusion of biometric data is illustrated in Figure 3.

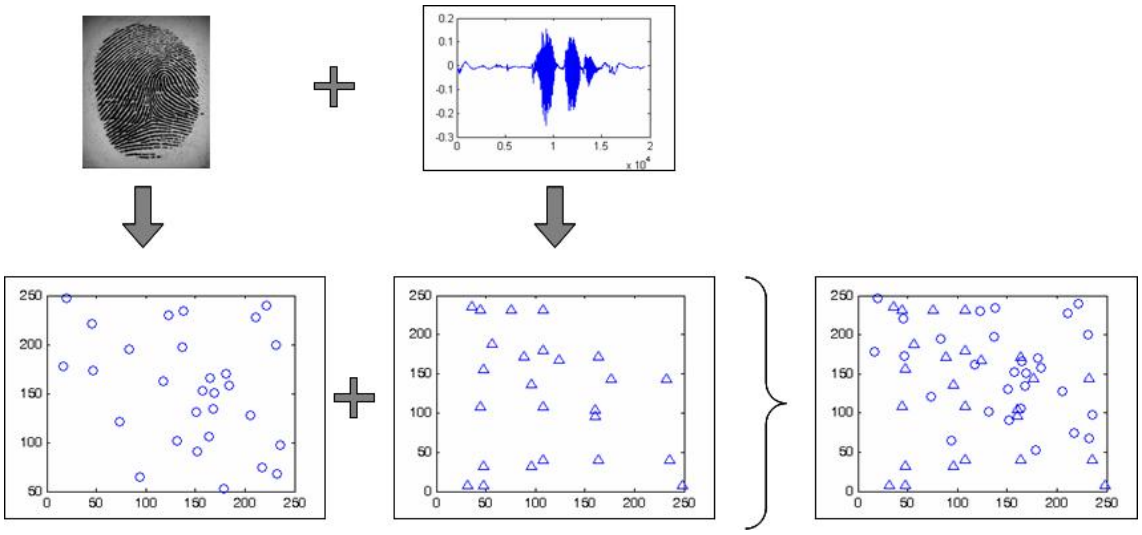

Figure 3. Template level fusion of biometric data: Minutiae points from the fingerprint and artificial points generated from voice are combined together in a user template. The points are marked as to indicate the source biometric, but this information is not stored in the database.

3.3. Matching

When a user comes for authentication, he or she is authenticated using both the fingerprint and voice modalities, in order. First, the fingerprint minutiae (A') are extracted and matched against the template (A+B) of the claimed user, consisting of both fingerprint and voice minutiae. The automatic matching is done via a simple matching algorithm finding the best alignment over all translations and rotations, allowing for some elastic deformation of the fingerprint (accepting two points as matching if they are within a small threshold in this alignment). After the alignment, the matched points are deleted from the template, leaving ideally only the points generated from the utterance of the claimed user (A+B-A'). Matching of the test fingerprint is illustrated in Figure 4.

Second, the test user provides the spoken password of the claimed user. Here, the utterance is processed and utterance minutiae (B') are generated as described in the previous section. Matching of the utterance minutiae with the remaining points in the template takes place in Hamming space, instead of Euclidean space to handle small noise in the utterance points. Small variations may result in unacceptable distances in the Euclidean space

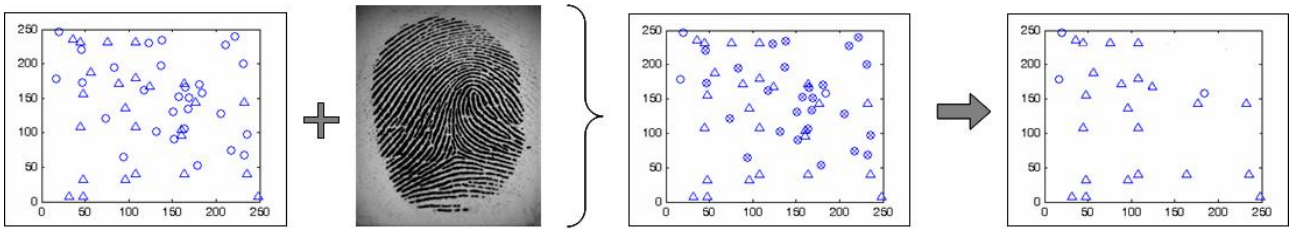

Figure 4. Illustration of the first phase of the verification, where the test fingerprint is matched with the user template shown on the left. Matched points of the template are marked with a cross.

depending on the position of the erroneous bit: if there is an erroneous bit in the most significant bit of the x or y coordinate, that utterance minutia will be far away from the reference utterance minutia in the Euclidean space, whereas it costs 1-bit error in the Hamming distance. To do this, all the remaining points in the template (A+B-A') are mapped back to the Hamming space by concatenating the x and y coordinates. Now all the remaining points are represented by 16 bit strings and the test utterance is also processed and divided into bit strings of length 16. To be considered a match, the Hamming distance between two bit strings should be less than or equal to a threshold (2-bit errors are accepted in this work). Then, matched utterance minutiae points are deleted from the template of the claimed identity. Matching of the test voice minutiae is illustrated in Figure 5.

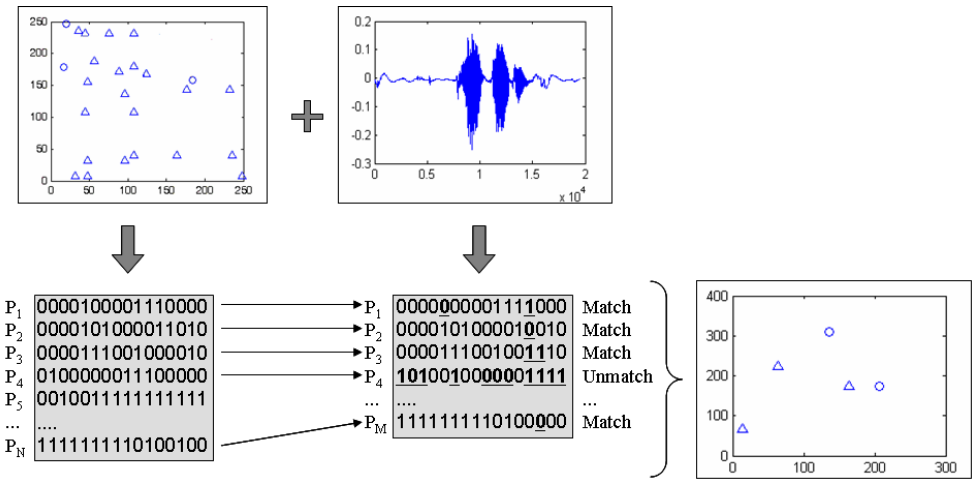

Figure 5. Illustration of the second phase of the verification where the utterance is matched with the remaining points in the user's template. Matched points are removed, showing here a successful match.

Finally, the decision for authentication is given according the score.[27] Note that the person is authenticated if a high percentage of the points involved in the second phase of the verification (remaining points from the template and the minutiae points of the query) is matched.

$$score = 2 \times \frac{|(A + B - A') \cap B'|}{|(A + B - A') + B'|} \tag{1}$$

In case A' matches A perfectly and B' matches B perfectly, the resulting score with this metric is 1, showing a good match. If A' was not successfully matched, it would be reflected in the final score since many minutiae points would be left unmatched, making the denominator large. If B' was not successfully matched, the numerator would be small.

4. DATABASE AND RESULTS

4.1. Database

The utterance database in this paper consisted of 30 users (15 men, 15 women). Each speaker is assigned a 6-digit password and asked to repeat their password 10 times. The passwords are grouped into 2-digit numbers (e.g. 35-45-67). After repeating his or her password, each user is asked to repeat the password of the previous user in the same gender. To close the loop, first user is asked to repeat the password of the last user in the database. This way, 10 genuine and 10 forged spoken passwords have been collected for 30 speakers.

The fingerprint database consists of 400 fingerprints from 200 individuals (2 imprints per person). We selected 30 users with good quality fingerprints and we matched each person with a randomly chosen person from the utterance database, to link these two unrelated databases. We then created 30 multi-biometric templates using one fingerprint of a person from the fingerprint database and the binary feature vector obtained from the 8 reference utterances of the matching person from the utterance database. In reference to the template, we will refer to these two people as the "user" from now on.

For genuine tests, we used the matching fingerprint and the 2 genuine utterances of the user. For forgery tests, we used the 199 non-matching fingerprints and the 10 utterance forgeries collected for the user. In other words, we are always testing the performance under the assumption that the password is known to an attacker.

4.2. Results and Future Work

The performance evaluation of the speaker verification system proposed in this paper is done by considering the false acceptance and false rejection rates of the system. The verification tests are done using 3 different scenarios: In the first scenario (FF), we tested the case where both fingerprints and utterance is a forgery. Hence, we had 2 genuine tests (1 fingerprint x 2 utterances) and 1990 forgeries (199 fingerprints x 10 utterances) for each user, generated by false fingerprints and false voice samples.

For the second scenario (FG), we tested the case where the fingerprint is a forgery, but the utterance is genuine. This may be the situation where the forger may have recorded a sample of the user whose voice is stored in the multi-biometric template. Hence, the EER is calculated by comparing the results from 2 genuine tests (1 fingerprint x 2 utterances) with 398 (199 fingerprints x 2 utterances) forgery tests for each user, generated by false fingerprint but genuine voice samples.

For the last scenario (GF), we tested the case where the fingerprint is genuine, but the utterance is forgery. This may be the situation where the forger may have acquired a silicon fingerprint of the user whose biometric is stored in the multi-biometric template. Hence, the EER is calculated by comparing the results from 2 genuine tests (1 fingerprint x 2 utterances) with 10 (1 fingerprint x 10 utterances) forgery tests for each user, generated by genuine fingerprint but false utterances. The results for this work can be seen in Table 1 for males and females separately.

	FF	**FG**	**GF**
females	0,54	3,50	16,60
males	1,01	6,60	26,00
mean	**0,77**	**5,50**	**21,30**

Table 1. Equal Error Rates: Equal error rates are given for three scenarios which indicates three different forgery attack types. These are: false fingerprint, false voice (FF); false fingerprint, genuine voice (FG) and genuine fingerprint, false voice (GF).

As can be see in Table 1, in a forgery attempt with a false fingerprint and false utterance, the equal error rate is only 0.77%. This rate increases to 5.5% when the attacker has access to the utterance of the user. Notice that the case when the attacker has access to the fingerprint of the user, the error rate significantly increases; however, even then about 80% of the attacks are repelled.

For comparison, in our previous work for unimodal voice verification,[5] EER was 3,3% with the same voice feature extraction algorithm for the same speaker database. Hence, average results obtained for impostors with mismatched fingerprint and voice (0.77%), are significantly improved over the unimodal system, as expected.

Furthermore, we have experimented with different quantization level and different partitioning of the data, as well as using PCA for projecting the 6-dimensional parts of the feature vector into 2-dimension (x,y) with similar initial results. We are experimenting with them further to obtain more reliable 2D features from voice. For future work, we also plan to measure how the multi-biometric template scheme preserves privacy, by calculating recall and precision rates for retrieving templates given only one of the biometric modalities, as previously done by Yanikoglu and Kholmatov.[27]

5. SUMMARY

The main goal of this work is to propose a multi-modal biometric system which preserves privacy and increases accuracy. Privacy is preserved by fusing biometric information from fingerprint and voice at the template level in the minutiae space. Since fingerprint minutiae and artificial points generated from voice are combined in the users' templates, privacy concerns are diminished by hiding the nature of the points in the templates. Furthermore, results from the tests show that the error rates are lower than using only voice as the unimodal biometric when compared with our previous work.

6. ACKNOWLEDGMENTS

The work presented in this paper has been supported by the project of The Scientific and Technological Council of Turkey with project number 105E165.

REFERENCES

1. Amengual J. C., Juan A., Prez J. C., Prat F., Sez S. and Vilar J. M. Real-time Minutiae Extraction in Fingerprint Images. *Proc. of the* 6^{th} *Int. Conf. on Image Processing and its Applications*, July 1997.
2. J. R. Bellegarda, D. Naik, M. Neeracher, K. E. A. Silverman. Language-independent, Short-enrollment Voice Verification over a Far-field Microphone, *IEEE International Conference on Acoustics, Speech and Signal Processing*, 1:445-448, 2001.
3. S. Ben-Yacoub, Y. Abdeljaoued, and E. Mayoraz. Fusion of Face and Speech Data for Person Identity Verification. *IEEE Trans. Neural Networks*, vol. 10, no.5, pp. 1065-1075, 1999.
4. R. Brunelli and D. Falavigna. Person Identification Using Multiple Cues. *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 17, no. 10, pp. 955-966, Oct. 1995.
5. E. Camlikaya, B. Yanikoglu, H. Erdogan. Voice Verification Using HMM. *IS&T/SPIE Annual Symposium on Electronic Imaging*, January 2008.
6. Davida, G., Frankel, Y., Matt, B. On Enabling Secure Applications through On-line Biometric Identification. In: *Proc. of the IEEE 1998 Symp. on Security and Privacy*, Oakland, Ca. 148-157, 1998.
7. N. Fan, J. Rosca, and R. Balan. Speaker Verification with Combined Threshold, Identification Front-end, and UBM, *Fourth IEEE Workshop on Automatic Identification Advanced Technologies*, pp.112-117, Oct. 2005.
8. Y. He., J. Tian, X. Luo, T. Zhang. Image Enhancement and Minutiae Matching in Fingerprint Verification. *Pattern Recognition Letters*, vol.24, pp.1349-1360, October 2003.
9. L. Hong and A.K. Jain,. Integrating Faces and Fingerprints for Personal Identification. *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 2, p. 1295-1307, Dec. 1998.
10. A. K. Jain, L. Hong, R. Bolle. On-line Fingerprint Verification. *IEEE Transactions On Pattern Analysis and Machine Intelligence*, Vol. 19, No. 4, pp. 302-314, 1997.
11. A. Jain, L. Hong, S. Pankanti, R. Bolle. An Identity Authentication System Using Fingerprints. *Proc. of the IEEE*, vol.85, pp. 1365-1388, September 1997.
12. X. Jiang, W. Yau. Fingerprint Minutiae Matching Based on the Local and Global Structures. *Proceedings of the 15th International Conference on Pattern Recognition*, Bacelona, vol. 2, pp.1038-1041, September 2000.
13. Juels, A., Sudan, M. A Fuzzy Vault Scheme. In:*Proc. of the 2002 IEEE Int. Symp. on Inf. Theory*, Lausanne, Switzerland 408, 2002.

14. J. Kittler, M. Hatef, R.P.W. Duin, and J. Matas. On Combining Classifiers. *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 226- 239, Mar. 1998.
15. Q. Li, B.-H. Juang, C.-H.Lee, Q.Zhou, and F.K. Soong. On Speaker Authentication, *IEEE Workshop on Automatic Identification Advanced Technologies*, Stony Brook, NY, pp.3 - 6, Nov. 1997.
16. Linnartz, J.P., Tuyls, P. New Shielding Functions to Enhance Privacy and Prevent Misuse of Biometric Templates. In: *Proc. of the 4th Int. Conf. on Audio and Video Based Biometric Person Authentication*, Guildford, UK 393-402, 2003.
17. D. Maltoni, Dario Maio, Anil K. Jain, Salil Prabhakar. Handbook of Fingerprint Recognition. Springer, New York, 2003.
18. F. Monrose, M. K. Reiter, Q. Li and S. Wetzel. Cryptographic Key Generation From Voice. In *Proceedings of the 2001 IEEE Symposium on Security and Privacy*, May 2001.
19. N.K. Ratha, V.D. Pandit. Robust Fingerprint Authentication Using Local Structural Similarity. *Fifth IEEE workshop on applications of computer vision*, Palm Springs, CA, pp. 29-34. December 2000.
20. Ratha, N., Connell, J., Bolle, R. Enhancing Security and Privacy in Biometrics-Based Authentication Systems. *IBM Systems Journal*, 40 614-634, 2001.
21. N. Ratha, S. Chen, A. Jain. Adaptive Flow Orientation Based Feature Extraction in Fingerprint Images. *Pattern Recognition 28*, pp. 1657-1672, 1995
22. J. G. Rodrguez, S. Cruz and J. Ortega. Biometric Identification through Speaker Verification over Telephone Lines, *Proceedings of IEEE Carnahan Conference on Security Technology*, pp. 238-242, ISBN: 0-7803-5247-5, Madrid, 1999.
23. A. Ross, A. Jain, A.K. Mutlimodal Biometrics: an Overview. *Proceeding of European Signal Processing Converence*, pp. 1221-1224, Austria, September 2004.
24. R. Snelick, U. Uludag, A. Mink, M. Indovina A. Jain. Large-Scale Evaluation of Multimodal Biometric Authetication Using State-of-the-Art Systems. *IEEE Trans. on PAMI*, vol:27, no:3, March 2005.
25. M. Tico, P. Kuosmanen. Fingerprint Matching Using an Orientation-Based Minutia Descriptor. *IEEE Trans. PAMI*, vol.25, no. 8 , pp. 1009-1014, August 2003.
26. M. Tico, P. Kuosmanen. An Algorithm for Fingerprint Image Post-processing. *Proc. of the* 34^{th} *Asilomar Conference on Signals, Systems and Computers*, vol.2, pp. 1735-1739, November 2000.
27. B. Yanikoglu, A. Kholmatov. Combining Multiple Biometrics to Protect Privacy. *Proceedings of ICPR-BCTP Workshop*, Cambridge, England, August 2004.
28. S. Young, J. Jansen, J. Odell, D. Ollason, and P. Woodland. *The HTK Book* (for HTK Version 2.1), Cambridge University Press, Cambridge, 1997.