International Journal of Advanced Trends in Computer Applications
www.ijatca.com

Feature-Level Fusion of Speech, Signature and Tongue using ordered weighted average using GA

[1]Harmanpreet Kaur, [2]Gaganpreet Kaur
[1]Research Fellow, [2]Asst Professor
[1,2]Sri Guru Granth Sahib World University, Fatehgarh Sahib, Punjab.
[1]har.innocent@yahoo.com, [2]gagan1492003@yahoo.com

Abstract: *In the proposed work, three modalities speech, signature and tongue are fused together. Features of tongue, speech and signature are extracted using SIFT, MFCC and DCT respectively. Fusion of three modalities is performed using ordered weighted averaging with modified GA. Features for non noisy and noisy samples have been collected separately. Proposed system also works efficiently on filtered noisy modalities. Motion blur filter is used for filtration of noisy samples.*

Keywords: *Biometric, multimodal, feature, SIFT, MFCC, DCT, GA, fitness function, fusion, ordered weighted averaging, motion blur filter.*

1. Introduction

Biometric systems are authentication systems which recognize a person's identity from the feature vectors extracted from his/her physiological and/or behavioural traits and matching it with information stored in database [1]. A biometric system that is based on extracting physiological characteristics are more reliable than systems using behavioural characteristics. Some of physiological characteristics that can be used as biometric are fingerprints, hand geometry, retina, iris, face, DNA, ear, signature, speech.

2. Tongue Biometric

The tongue is a unique organ to each person in its shape and texture. Tongue can be stuck out of mouth for inspection and is the only inner organ that can be used in biometric authentication [2]. Moreover tongue cannot be forged as it is protected inside the mouth.

Fig. 2 *Different shapes of tongue.*

In proposed work texture features of tongue are calculated by SIFT (Scale invariant feature transform) algorithm.

Fig.3 *Different textures of tongue.*

2.1 Database for tongue

For acquiring tongue images, a digital camera is used. The height and distance of digital camera is taken into consideration to ensure good clarity and accuracy of tongue images being taken. The height of camera is supported by a tripod stand of 14 cm and person's head is supported by head rest of height 15 cm. The distance between head rest and camera is 9 cm [3]. The images of tongue should be taken in an enclose area to ensure lighting of tongue image in control.

In proposed work, 11 images of tongue of 11 different people's have been collected and stored in database.

Fig. 4 *Database of tongue.*

2.2 SIFT Algorithm

SIFT (Scale invariant feature transform) algorithm is a feature vector generation algorithm which takes an image as input and generates a large collection of local feature vectors [2]. SIFT algorithm extracts features by applying four stage filtering approach:

i. **Scale-space extrema detection:** Difference of Gaussians is applied for locating scale-space extrema.

ii. **Keypoint localization:** This is achieved by calculating the Laplacian value for each keypoint found in (i). If the function value is below a threshold value then this point is excluded. This removes extrema with low contrast.

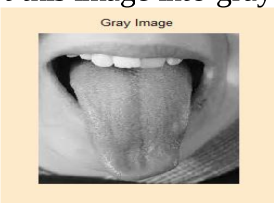
iii. **Orientation assignment:** One or more orientations are assigned to the keypoints based on local image properties.

iv. **Keypoint Descriptor:** The local gradient data is used to measure keypoint descriptors. Then these are transformed and a set of histograms are created over a window centred over keypoint.

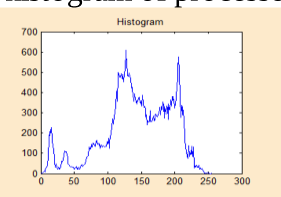
2.3 Implementation Steps

Step 1: Collect tongue image that can be easily processed

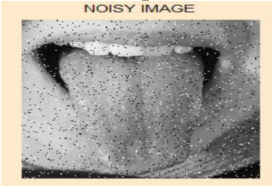
Step 2: Convert this image into gray scale image



Step 3: Create histogram of processed image

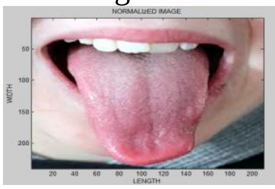


Step 4: Add noise in the uploaded image

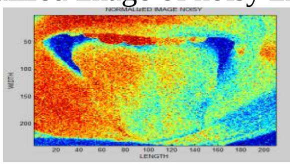


Step 5: Apply motion blur filtering to the noisy image

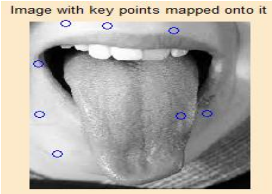
Step 6: Normalized image of base image



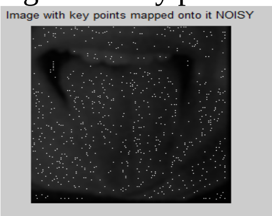
Step 7: Normalized image of noisy image



Step 8: Apply SIFT algorithm to find features in the image



(a) Base image with key points mapped onto it



(b) Noisy image with key points mapped onto it

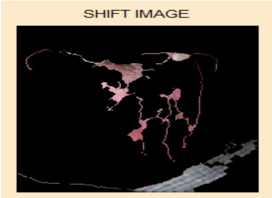
Step 9: Edge image is created

(a) Edge image for non noisy

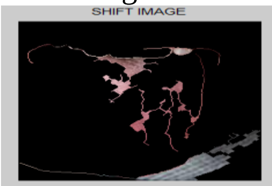
(b) Edge image for noisy

ISSN: 2395-3519

Step 8: SIFT image



(a) SIFT image for non noisy



(b) SIFT image for noisy

3. Signature Biometric

Signature is a behavioural biometric. Signature recognition system can be offline or online. Online signature is easy but it requires special devices for data acquisition [4]. Signature is a possible database that can be very easily collected anywhere anytime. In the proposed work features of signature are extracted using DCT algorithm.

3.1 Database

Database used in for the proposed work is SVC20. SVC20 is the subset of SVC2004 database. 20 original and 20 forged signatures for 20 signers are there in database.

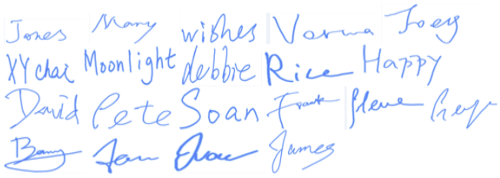


Fig. 5 *Database for signature.*

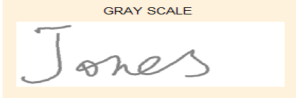
3.2 DCT (Discrete Cosine Transformation)

The discrete cosine transform (DCT) is closely related to the discrete Fourier transform. It is a separable linear transformation; that is, the two-dimensional transform is equivalent to a one-dimensional DCT performed along a single dimension followed by a one-dimensional DCT in the other dimension.

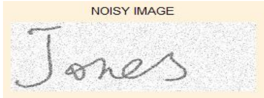
3.3 Implementation steps

Step 1: Upload signature image.

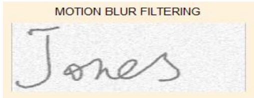
Step 2: Convert this image into gray scale image.



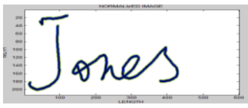
Step 3: Add noise to the image to check the performance of the system.



Step 4: Apply motion blur filtering to noisy image



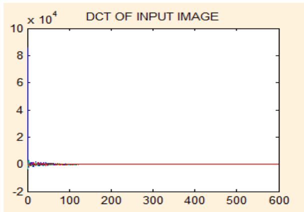
Step 5: Normalization is performed.



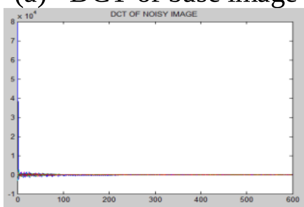
(a) Normalized base image

(b) Normalized noisy image

Step 6: Apply DCT algorithm.

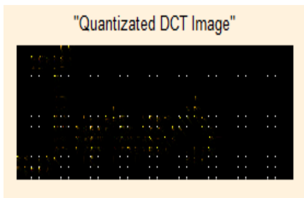


(a) DCT of base image

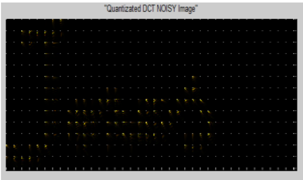


(b) DCT of noisy image

Step 7: Find quantized DCT image.



(a) Quantized image of base image



(b) Quantized image of noisy image

4. Speech Biometric

Speech biometric is commonly used as authentication system. Speech recognition is identifying speakers form a group of speakers using their speech to identify their identity. Database for speech is easy to collect. In the proposed work features of speech are extracted using MFCC algorithm.

4.1 Database

The CMU_ARCTIC databases were constructed at the Language Technologies Institute at Carnegie Mellon University as phonetically balanced, US English single speaker databases designed for unit selection speech synthesis research. The databases consist of around 1150 utterances carefully selected from out-of-copyright texts from Project Gutenberg. The databases include US English male and female speakers. In the proposed work only male database has been used which contains 1132 utterances spoken by a US English male speaker which was recorded at 16bit 32KHz, in a sound proof room, in stereo, one channel was the waveform, the other EGG. The database was automatically labelled using EHMM an HMM labeller that is included in the Fest Vox distribution. No hand correction has been made.

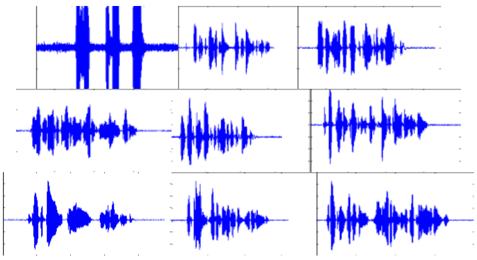


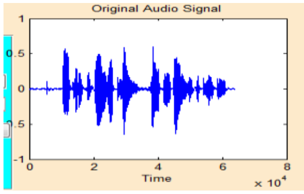
***Fig. 6** Samples from database*

4.2 MFCC Algorithm

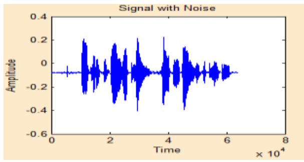
- Take the Fourier transform of a signal.
- Map the powers of the spectrum obtained above onto the mel scale.
- Take the logs of the powers at each of the mel frequencies.
- Take the discrete cosine transform of the list of mel log powers, as if it were a signal.
- The MFCCs are the amplitudes of the resulting spectrum.

4.3 Implementation Steps

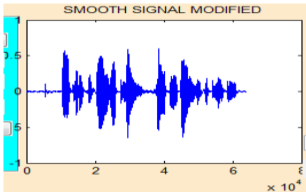
Step 1: Upload input signal.



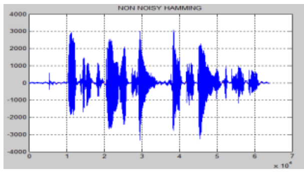
Step 2: Add noise to check the performance of the system.



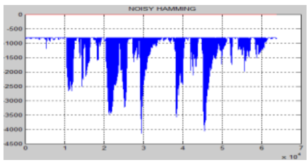
Step 3: Smooth (filtered) signal modified.



Step 4: Hamming window

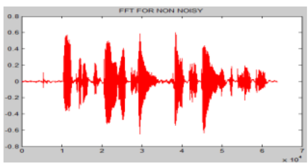


(a) Hamming window for non noisy

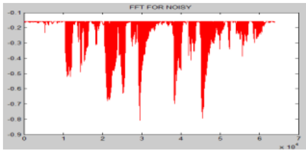


(b) Hamming window for noisy

Step: 5 FFT step for noisy and non noisy

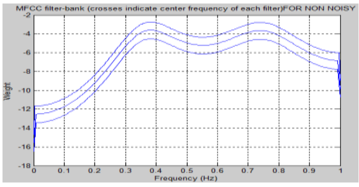


(a) FFT step for non noisy

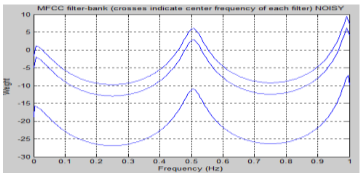


(b) FFT step for noisy

Step 6: Apply MFCC to find MFCC filter-bank

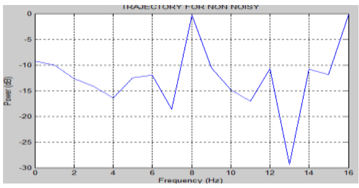


(a) MFCC filter bank for non noisy

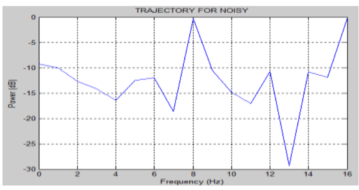


(b) MFCC filter bank for noisy

Step 7: Find trajectory graph

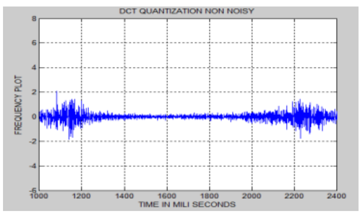


(a) Trajectory for non noisy

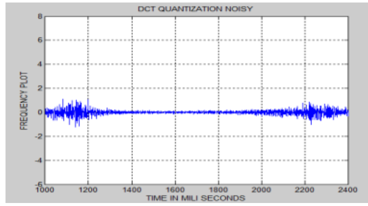


(b) Trajectory for noisy

Step 8: Find DCT quantization.



(a) DCT quantization for non noisy



(b) DCT quantization for noisy

5 . FUSION

1. Load sample (noisy or non-noisy).
2. Applying concept of apriori two itemsets are calculated.
3. Support is calculated for all itemsets.
   Supp(X U Y)=count(X U Y) /count(N) [13].
4. Those two item sets with support greater than minimum support are stored in database.
5. Call GA.

Modified Genetic algorithm

1. Population size= 500 //default it is 20 which is not suitable for larger problems
2. Mutation function=0.1 // default it is 0.01
3. Crossover=0.95 // default it is 0.80
4. Fitness function of GA is enhanced using ordered weighted averaging and concept of threshold.
   threshold_wt=w1+w2+w3
   threshold_new=threshold_wt/3
5. Check if the new threshold is less than the previous threshold for the current set, threshold is ignored else accepted.

6. Proposed Methodology

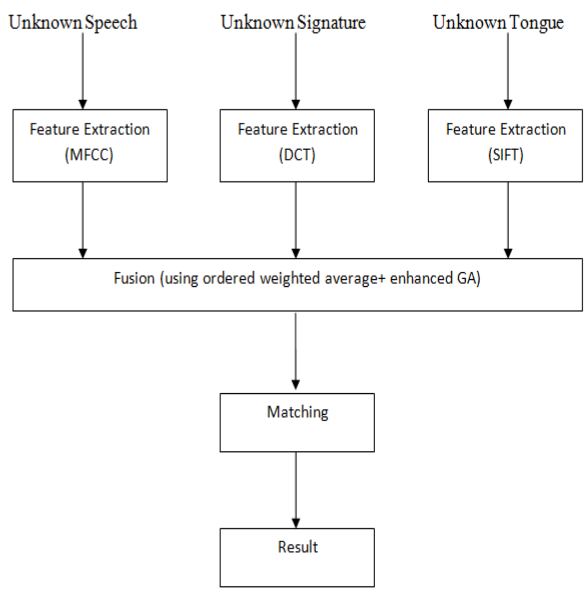
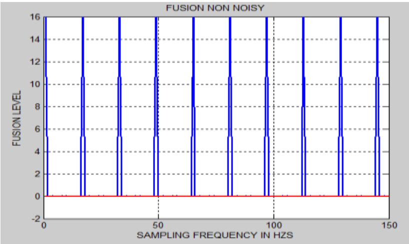


Fig. 7 *Flowchart for proposed work*

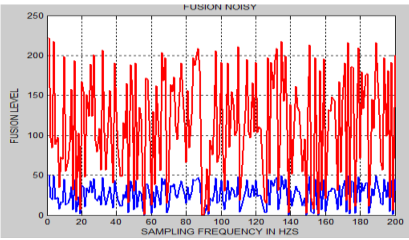
7. Results

In the proposed work, fusion is done for samples without noise and samples with noise and accuracy is checked for both fusions and also total time for both fusions have been drawn using graphs.

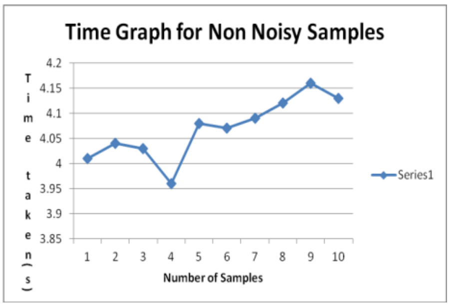
(A) Fusion for non noisy samples



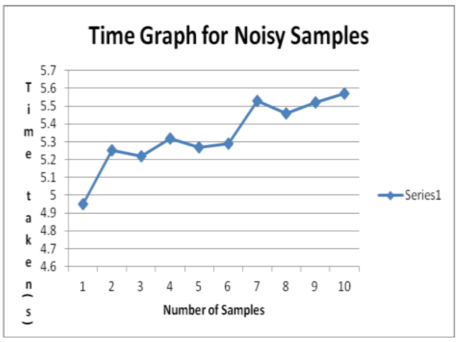
(B) Fusion for noisy samples



(C) Time graph for non noisy samples (time in seconds)



(D) Time graph for noisy samples (time in seconds)



(E) FAR, FRR and Accuracy

FAR= Number of falsely accepted samples/Total number of samples
FRR= Number of falsely rejected samples/Total number of samples

For non-noisy samples

FAR= 4/80 =0.05
FAR=0.05
FRR= 4/80=0.05
FRR=0.05
Accuracy= (1-(FAR+FRR))
Accuracy= (1-(0.05+0.05))
Accuracy=1-.10=.90=90%

For noisy samples

FAR=4/80=0.05
FAR=0.05
FRR=12/80=0.15
FRR=0.15
Accuracy= (1-(FAR+FRR))
Accuracy= (1-(0.05+0.15))
Accuracy=1-0.2=.80=80%

8. Future Work

In future, new feature extraction algorithms can be used to extract features of various modalities. Work can be done to improve the accuracy of the system and to reduce the time taken by the system. Also work can be done by using other types of noise other than Gaussian and can be removed using different types of filters.

References

[1] A.K. Jain, A. Ross, and S. Prabhakar, "An Introduction to Biometric Recognition", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 1782-1793,2004.

[2] M. Diwakar, and M. Maharishi, "An Extraction and Recognition of Tongue-Print Images for Biometrics Authentication System", *International Journal of Computer Applications*, vol. 61, no. 3, pp. 36-42, 2013.

[3] Z. Liu, J.Q. Yan , D. Zhang, and Q.L. Tang, "A Tongue-Print Image Database For Recognition", in *Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, Hong Kong*, pp. 19-22,2007.

[4] P. Karthik, R.V.S.S. Prasad, and S.R.M. Prasanna, "Noise Robust Multimodal Biometric Person Authentication System using Face, Speech and Signature Features" , in

India Conference, 2008, INDICON 2008, Annual IEEE, pp.23-27, 2008.

[5] A.S. Frisch, A. Riera, and S. Dune, "Fusion operators for multi-modal biometric authenticiation based on physiological signals", in *Fuzzy Systems(FUZZ),2010IEEE International Conference*, pp.1-7,2010.

[6] K. Delac, and M. Grgic, "A Survey of Biometric Recognition Methods", in *46th International Symposium Electronics in Marine, ELMAR-2004*, pp. 184-193, 2004.

[7] D. Kaur, G. Kaur, and D. Singh, "Efficient and Robust Multimodal Biometric System for Feature Level Fusion (Speech and Signature)", *International Journal Of Computer Applications*, vol. 75, no. 5, pp. 33-38, 2013.

[8] V. Conti, C. MiliTello, F. Sorbello, and S. Vitabile, "A Frequency-based Approach for Features Fusion in Fingerprint and Iris Multimodal Biometric Identification Systems", *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews*, vol. 40, no.4, pp. 384-395, 2010.

[9] M.F. Zanuy, "Data Fusion in Biometrics", *IEEE A & E Systems Magazine*, pp.34-38, 2005.

[10] D. Kaur, and G. Kaur, "Level of Fusion in Multimodal Biometrics: a Review", *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, issue 2, pp. 242-246, 2013.

[11] . Ross, and A. Jain, "Information Fusion in Biometrics", *Pattern Recognition Letters*, vol. 24, no. 13, pp. 2115-2125, 2003.

[12] S. Prabhakar, and A. Jain, "Decision-level in Fingerprint Verification", *Pattern Recognition*, vol. 35, no. 4, pp. 861-874, 2002.

[13] S. Sharma, and V. Chopra, "Association Rule Mining: A Multi-objective Genetic Algorithm Approach Using Pittsburgh Technique", in *International Journal of Recent technology and engineering*, vol. 2, issue 4, pp.121-124, 2013.

[14] S.B. Yacoub, Y. Abdeljaoued, and E. Mayoraz, "Fusion of Face and Speech data for Person Identity Verification", in *IEEE Transactions on Neural Networks*, vol. 10, no. 5, 1999.

[15] A. Rattani, D.R. Kisku, M. Bicego, and M. Tistarelli, "Feature Level Fusion of Face and Fingerprint Biometrics", in *Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007. First IEEE International Conference* , pp. 1-6, 2007.

[16] A. C. Ramachandra, S. K. Abhilash, K.B. Raja, K. R. Venugopal, and L. M. Patnayak, "Feature Level Fusion Based Bimodal Biometric Using Transformation Domine Techniques", in *IOSR Journal of Computer Engineering (IOSRJCE)*,vol. 3, issue 3, pp. 39-46, 2012.

[17] Y. Meng, and Dr. B. Tiddeman, "Implementing the Scale Invariant Feature Transform (SIFT) Method".