Nallagatla, Vishnu P. and Chandran, Vinod (2010) ***Sequential decision fusion for controlled detection errors.*** In: 13th International Conference on Information Fusion, 26-29 July 2010, Edinburgh International Conference Centre, Edinburgh.

Sequential Decision Fusion for Controlled Detection Errors

Vishnu Priya Nallagatla
School of Engineering Systems
Queensland University of Technology
Brisbane, Australia.
v.nallagatla@qut.edu.au

Vinod Chandran
School of Engineering Systems
Queensland University of Technology
Brisbane, Australia.
v.chandran@qut.edu.au

Abstract - Information fusion in biometrics has received considerable attention. The architecture proposed here is based on the sequential integration of multi-instance and multi-sample fusion schemes. This method is analytically shown to improve the performance and allow a controlled trade-off between false alarms and false rejects when the classifier decisions are statistically independent. Equations developed for detection error rates are experimentally evaluated by considering the proposed architecture for text dependent speaker verification using Hidden Markov Model (HMM) based digit dependent speaker models. The tuning of parameters, n classifiers and m attempts/samples, is investigated and the resultant detection error trade-off performance is evaluated on individual digits. Results show that performance improvement can be achieved even for weaker classifiers (FRR-19.6%, FAR-16.7%). The architectures investigated apply to speaker verification from spoken digit strings such as credit card numbers in telephone or VOIP or internet based applications.

Keywords: Multi-instance fusion, multi-sample fusion, detection error trade-off, sequential decision fusion

1 Introduction

The major concern in a biometric verification system is its accuracy. One general problem of biometric system is that the individual samples of the same person are not identical for each presentation. This intra-class variability is caused by several reasons such as different environments, changing sensors or even natural biometric variability. Inter-class similarity is achieved by high degree of identicalness of the same biometric trait between different persons. These limitations may lead to misclassification of the verification claims resulting in false alarms and false rejects. These two errors are dependent and in general it is difficult to reduce one type of error without increasing the other. The main focus of this paper is to obtain better trade-off between both the detection errors using information fusion techniques.

In the context of biometrics, information fusion refers to the use of multiple sources of biometric information to obtain a decision. Such systems, known as multi-biometric systems, can improve the accuracy of a biometric system. Based on the nature of information sources being consolidated, multi-biometric systems can be classified into 6 categories [1]: multi-sensor, multi-algorithm, multi-instance, multi-sample, multimodal and hybrid. Jain et al. performed experiments on fingerprint system and have show that multi-instance (two fingers) or multi-sample (two impressions of the same finger) fusion results in improved performance [2]. This paper presents architecture of multi-biometric system that integrates the multi-instance [3] and multi-sample [4] fusion schemes for controlling the trade-off between the detection error rates.

The architectures used for integrating the fusion schemes could be either serial or parallel [1]. The use of particular type of architecture is mainly application dependent. The serial architecture is considered for the acquisition and processing of information in this paper. The reason/motivation for choosing this architecture is explained in the section 2. In a serial approach, the acquisition and processing of biometric samples takes place sequentially and so the decision outcome from one biometric system may affect the processing of the subsequent systems [1].

In this paper, the scenarios in which the integration of multi-instance and multi-sample fusion schemes can be applied are explained in section 2. In section 3, the framework for the multi-instance, multi-sample and the proposed fusion methods is presented with theoretical prediction of detection error rates. The frame work is explained in the context of text-dependent speaker verification system. The methodology used for performance evaluation and the results obtained are explained in the section 4 and finally, in Section 5, a brief conclusion and the possible future work are presented

2. Application Scenario of Multi-instance and Multi-sample Fusion Schemes

Most commercial applications of biometrics (for example, telephone banking, access control or e-commerce) includes bi-factorial authentication (combination of knowledge and biometrics). The client/user in these applications presents the biometric information of some specific knowledge (identification PIN/ credit card number/ password) to the verification system. The biometric characteristics extracted for verification can be either the user's uniqueness in uttering the knowledge information, his writing style or even the way he types the information. The identity claim in this application scenario can be verified by classifying the entire knowledge information at once (single instance) or by fusing classification information from individual digits/characters of the knowledge information (multiple instances).

With a multi-instance system, each digit/character is processed sequentially using a different classifier and so each instance has the ability to independently produce a decision about the user's claim. In this approach, a rejection at any one of the classifier in the sequence results in a final decision of rejecting the identity claim. This fusion method efficiently reduces the false acceptances as it is hard for an impostor to reproduce a true user's characteristics for multiple instances. However, there is also a possibility for a true user to get wrongly rejected at any stage of classification because of a large intra-class variation. This increases the number of false rejections. This approach is well suited for high security application scenarios, e.g., logging in as super-user where providing access to unauthorized individuals is to be restricted to a minimum possible. However, this method is not desirable in most of the banking and point of service applications where a low false rejection rate causes greater customer convenience.

In a traditional password based systems, the user is allowed with certain number of attempts/tries (usually 3 attempts) to get verified by the system. Similar approach can be adopted in multi-biometric systems by considering repetition of samples from the same biometric characteristic. This method of multi-sample fusion helps in reducing the genuine user rejections but increases the false acceptances as the impostor is given additional chances for verification. Restricting the number of multiple samples to a minimum can limit increase in FAR to certain extent. This is because, in practice, a false claimant/impostor usually requires more number of attempts to get accepted rather than a true user who will be good in adapting the biometric characteristics to his/her own model.

It can be noted that the multi-instance and multi-sample fusion schemes reduce one type of error at the cost of increase in the other detection error. So this paper presents an architecture that considers the integration of both multi-biometric fusion schemes to arbitrarily reduce both the errors. The performance of the proposed architecture is evaluated by verifying a user based on his unique speech characteristics (speaker verification).

Typical applications of the proposed architecture using speaker verification includes telephone and internet banking, information services, security control, remote access to computers, telephone and internet based shopping, etc. However it is desirable in most of these applications to set the parameters, number of samples/attempts and the number of instances, to be used for verification of a specific speaker before performing real-world verification. This paper presents formulae in the next section that can be used to tune these parameters.

3. Multi-biometric fusion for speaker verification

Speaker verification is a process of making a decision to either accept or reject the identity claim of a speaker. The basic structure for a speaker verification system is explained in [5]. The verification decision is usually based on a likelihood score obtained by comparing the test utterance to the claimant's model. The most commonly used technique to model a claimant in text-dependent speaker verification system is the HMM [6].

3.1 Framework of the multi-instance fusion system

An instance in the context of speech refers to text spoken by an individual, when modelled, has the ability to discriminate the speaker from others. In a text-dependent mode, multiple speaker specific models can be trained by varying the text (words or phrases).

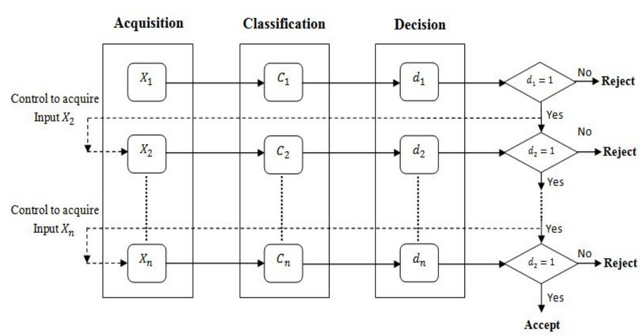


Figure 1. Architecture for a multi-instance/ multi-classifier fusion scheme with '*n*' classifiers arranged sequentially

The architecture of multi-instance system is shown in Fig. 1. There is a sequential chain of classifiers $C_1, C_2, C_3, \ldots C_n$ with each classifier verifying an input test utterance $X_1, X_2, X_3 \ldots X_n$ respectively. The classifier C_i in this context refers to an

HMM, modelled using the training data of the instance '*i*'. Whenever classifier C_{i-1} accepts the input data X_{i-1}, the control is given to acquire the input for next classifier in the sequence, C_i ($2 \leq i \leq n$). This is similar to the application of AND logic where the final decision (d) of the system is to accept (d=1) the claim only if the decisions from individual classifiers (d_1=1, d_2=1, d_3=1 .... d_n=1) is to accept the speaker.

Each decision d_i of a classifier is characterized by two error probabilities: the probability of a false acceptance, α and the probability of false rejection rate, ρ. Considering the decisions d_i, i = 1, 2, ...n from each of the classifier to be statistically independent, the application of AND Rule can be used for fusing the decisions. The False Acceptance Rate (FAR) for the fused system is

$$\alpha_{Comb} = \prod_{i=1}^{n} \alpha_i \quad (1)$$

To analyze the AND rule it is more convenient to work with the detection probability, p_d = 1- ρ. The detection rate for the fused decision is given by

$$p_d(\alpha_{Comb}) = \prod_{i=1}^{n} p_{d,i}(\alpha_i) \quad (2)$$

Considering the false acceptance rate of each classifier be α and the false rejection rate be ρ, the resulting FAR is given as

$$\alpha_{Comb} = \alpha^n \quad (3)$$

Converting equation (2) into terms of False Rejection Rate (FRR)

$$\rho_{Comb} = \rho + (1-\rho)\rho + (1-\rho)^2\rho + ...(1-\rho)^{n-1}\rho$$
$$\rho_{Comb} \approx n\rho \ \text{(when } \rho << 1\text{)} \quad (4)$$

It can be noted that the reduction in the false acceptance rate is multiplicative (Equation 3) while the increase in the false rejection rate is approximately additive (Equation 4) which is desirable in most of the high security applications. The assumption of statistically independent decisions here is an ideal one but by using speaker dependent HMM classifiers for each instance, an assumption of independence is likely to be good when the phonemes involved in the word are different and will hold reasonably well even when they share some phonemes but differ in the order in which they are put together.

3.2 Framework of multi-sample fusion system

The architecture of multi-sample system presented in the section 2 is shown in Fig. 2. This architecture is similar to the method proposed by Nelson and Kashi [7] on signature verification system. For this architecture, the maximum allowed number of repeated samples, *m*, need to be fixed prior based on the error rates obtained from a single sample system. In a multi-sample system, the speaker presents an input test utterance X_i *(i=1, 2, ...m)* and the classifier *C* makes a decision to either accept or reject the speaker.

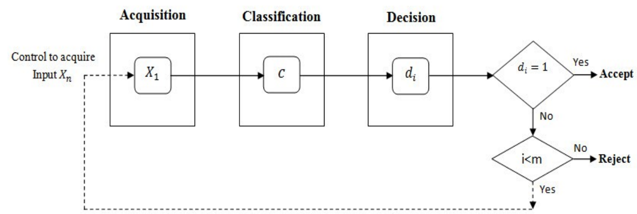


Figure 2. Architecture for multi-sample fusion with '*m*' repetition of samples

If the claim is accepted (d_i=1), the system does not go for another sample and the speaker is declared to be genuine. If the claim is rejected the speaker is allowed to present a repeated sample (X_{i+1}) of the same text. The number of multiple samples/attempts ('*i*') initially being 1 adds up with every successive attempt and can be repeated until either the speaker is accepted or the number of repeated attempts reaches the maximum allowed ('*m*'). In case the speaker fails to get verified within the maximum allowed attempts, the claim is rejected.

For a speaker to be declared genuine, it is sufficient if any one sample presented to the system gets accepted and so an OR logic can be used for acceptance. However, the speaker is considered to be an impostor when all the '*m*' repeated samples are rejected and so AND logic is used for rejection. Considering the probability of false acceptance and false rejection for each independent tries to be α and ρ respectively, the FAR and FRR for the fusion scheme can be given by:

$$\alpha^{Comb} = m\alpha \quad (5)$$
$$\rho^{Comb} = \rho^m \quad (6)$$

From the Equations (5) and (6) it is clear that while the false rejection rate decreases (since α and ρ are less than 1), the false acceptance rate increases. In general, it would lead to conclusion that no significant gain could be achieved with multiple tries. However, the experiment conducted in [7] has shown that the FRR reduces significantly whereas the FAR increases only slightly.

It was shown in [8] that the control over the trade-off between errors achieved using cascaded multiple classifiers (Equations 3 and 4) get reversed for multiple attempts (Equations 5 and 6).

3.3 Framework of the proposed architecture

The proposed architecture is based on the integration of multi-instance and multi-sample fusion schemes. The integration is performed at each stage of classifier (instance) verification. The architecture can be explained based on decisions from the multi-instance and multi-sample systems:

a. If the classifier decision, for the sample of an instance, is to accept the speaker then the sample for the next instance in sequence is acquired and processed by a different classifier
b. If the classifier decision, for a sample of an instance, is to reject the speaker then a repetition of sample for the same instance is acquired and processed by the same classifier

The final decision of the proposed system for n number of classifiers (for '*n'*-instances) with each classifier allowing '*m'* number of multiple samples can be either to Accept or Reject the identity claim of the speaker. The final decision of the proposed system is to:

1. Accept - only if the speaker is accepted by all n classifiers in the sequence within the maximum number of allowed multiple attempts '*m*'.
2. Reject – if the speaker is not able to get accepted at any one of the classifier within the allowable number of multiple attempts '*m*'.

The detection error rates of the proposed system can be obtained by using the equations (3) and (5) for false acceptance rate and equations (4) and (6) for false rejection rate.

$$\alpha_n^m = (m\alpha)^n \quad (7)$$

$$\rho_n^m \approx n(\rho^m) \text{ (when } \rho << 1\text{)} \quad (8)$$

Assuming the response time for an instance verification to be *t/n* seconds, the trade-off on using '*m*' multiple presentations for '*n*' instances becomes the increase in total time for verification to an upper limit of '*mt*'[8]. However, the total verification time is often less than the upper limit. This is because, in general, the number of attempts required by a true speaker to get verified correctly is far less than that of an impostor. So there is a possibility for the true speaker to get accepted before reaching the maximum number of attempts and so the verification time at each instance is mostly less than '*mt*'. Further, in a sequential system, if the classifier decides to reject a speaker at any of the intermediate stage, the processing of samples for the subsequent instances does not take place. So in the case of a reasonably performing classifier, the total verification time for p number of instances with m attempts is less than '*mt*' (i.e., p*m*t/n < mt, p<n). Hence, it can be considered that the false acceptance rate can be reduced arbitrarily without trading off the false rejection rate, at the expense of some increased time for a verification process.

4. Experimental Setup

4.1 Database

In order to evaluate the performance of the proposed fusion scheme (multi-instance and multi-sample fusion) speech data related to multiple words/phrases with multiple repetitions for each word/phrase is needed. The database used for experiments in this paper is the CSLU [Centre for Spoken Language Understanding]: Speaker Recognition Version 1.1 database [9]. All of the data is collected over digital telephone lines and recorded using the CSLU T1 digital data collection system. The data recorded form each participant includes single words, phonetically rich sentences, digit strings, free speech, personal information and a mimicked sentence.

The experiments performed on the proposed system require multiple instances with repetition of data for each instance and so the digit strings from this database are used. The digit strings are sequences of 5 digits - P (5 3 8 2 4), Q (6 1 oh 9 7), R (4 0 7 1 3), S (2 8 3 7 6), T (1 9 oh 5 4) and U (0 5 2 3 9). The digit strings are segmented into individual digits manually for 11 speakers (randomly selected) and speaker models are created for the digits 1, 2, 3, 4, 5, 7 and 9. Digits 6 and 8 are discarded because of data insufficiency. For experiments performed, each digit is considered to be an instance and the repetitive sample is randomly picked from the remaining database.

4.2 Speaker Verification parameters

The performance of a speaker verification system largely depends on the parameters used at different stages. During feature extraction process, utterances are processed in 26 ms frames, Hamming windowed and pre-emphasized with a coefficient of 0.97. The feature set is formed by Mel-frequency cepstral coefficients (MFCC).

In training phase, Left - Right HMM models with five states per phoneme and three mixtures per state are created for each digit. A universal background model is used for speaker normalization and this model is adapted using Maximum a Posteriori (MAP) and Maximum Likelihood Linear Regression (MLLR). Client and background models have the same topology. In verification mode, impostor testing is done on the speech data from speakers other than the claimed identity. However, as it is a text-dependent system, the digit used as the input is matched to the corresponding claimed speaker model.

4.3 Results

The dataset is divided into train, tune and test subsets that are disjoint.

a) Train set: This set consists of 21 utterances for each digit used for training a digit dependent HMM model.
b) Tune set: The tune set for each digit has 35 utterances for genuine user testing and a total of 140 utterances (i.e., 14 utterances from each of 10 impostors) for impostor testing. This dataset is used for setting the thresholds and determining the Equal Error Rates for each individual digit.
c) Test Set: The performance of the proposed system is evaluated using the test set for different combinations of parameters in the classifier architecture. This set includes 70 utterances from genuine speaker and 420 utterances (i.e., 42 utterances for each of the 10 impostors) for testing the false acceptances of individual digits.

The error rates are obtained by performing speaker verification tests on 11 speakers, each time choosing one speaker as genuine and the other 10 speakers as impostors. The equal error rates for the tune set are evaluated by setting speaker dependent thresholds for each digit. These thresholds are used for determining the detection errors for each digit on the test set. The mean error rates for 11 speaker tests are presented in the table 1. The Equal Error Rate (ERR) from tune set is used to obtain the ideal error rates using the theoretical equations explained in the section 3. The test dataset is used to experimentally evaluate the theoretically obtained error rates.

Multi-instance fusion experiments:

The initial experiments are performed to evaluate the effectiveness of multi-instance fusion. The performance of the system is tested by progressively increasing the number of instances/digits used for verification. Figure 3(a) shows the error rates obtained for the multi-instance fusion method for 11 speakers. Each curves above and below the zero line represent the FRR and FAR respectively for each speaker obtained on different digit combinations.

Table 1. Error rates obtained on the tune and test datasets for individual digits shown with standard deviation

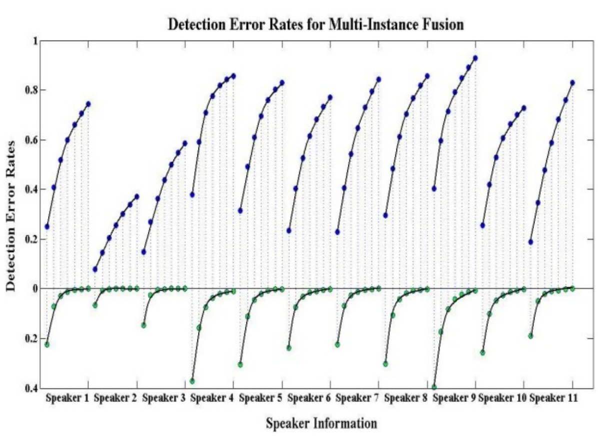
	Tune Set	Test Set	
Digits	EER	FRR	FAR
1	$0.2^{\pm0.11}$	$0.201^{\pm0.11}$	$0.2^{\pm0.11}$
2	$0.307^{\pm0.13}$	$0.308^{\pm0.13}$	$0.310^{\pm0.13}$
3	$0.216^{\pm0.13}$	$0.217^{\pm0.13}$	$0.216^{\pm0.13}$
4	$0.330^{\pm0.12}$	$0.339^{\pm0.12}$	$0.333^{\pm0.12}$
5	$0.281^{\pm0.09}$	$0.291^{\pm0.09}$	$0.281^{\pm0.09}$
7	$0.192^{\pm0.08}$	$0.195^{\pm0.08}$	$0.192^{\pm0.08}$
9	$0.208^{\pm0.12}$	$0.217^{\pm0.12}$	$0.207^{\pm0.12}$

There are 7 points on each curve each representing the number of digits used for verification. i.e., first point gives the mean error rates for tests on individual digits, second point represents the mean error rates for 2 digit combinations, third point is the mean error rate for tests on 3 digit combination and so on. The last data point on each curve is for the tests performed using all the digits (7 digit combination).

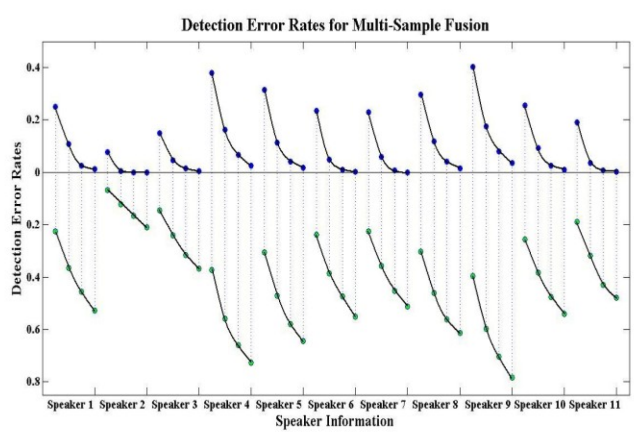
It is evident from the figure that multi-instance fusion results in lowering the number of false acceptances (curves below the 'zero' line) at the cost of increase in false rejection rate (curves below the 'zero' line). These results support the discussion in the section 3.1.

Multi-sample fusion experiments:

Figure 3(b) shows the FAR and FRR for multi-sample fusion method. The curves plotted are similar to the multi-instance fusion curve except that each point on the curve represents the number of repeated samples.



(a)



(b)

Figure 3(a). Plot for detection error rates of multi-instance fusion for 11 speakers **(b)** Plot for detection error rates of multi-sample fusion for 11 speakers (FAR curves – below the 'zero' line, FRR curves – above the 'zero' line)

The first point represents the tests performed on each digit without allowing any repetition of the samples. The second point is for the tests performed on digits with two multiple samples allowed and so on up to 4 data points. These experiments also support the discussion given for multi-sample fusion schemes in section 3.2. With the increase in number of multiple samples used, the number false rejection reduces where as the false acceptance rate increases.

Proposed multi-instance and multi-sample fusion experiments:

The proposed architecture is based on the integration of multi-instance and multi-sample fusion schemes (section 3.3). As noted from the figures 3 and 4, the performance improvement of the fusion schemes depends greatly on the individual digit classifier performance. So the analysis of the proposed method is carried on by selecting two speakers whose performance is good (speaker 2) and worse (speaker 9) compared to other speakers (as observed from the figures 3 and 4).

Figure 5(a) and 5(b) presents the detection error rates for the speaker 2 and speaker 9 respectively. The figure shows the mean error rates obtained by tuning the parameters n (number of classifiers) and m (number of attempts). The curves in the figure represent the error rates for the use of multiple samples and the seven points on each curve represent the digit combinations increasing progressively from bottom right of the figure to the top left. The points below the line for the data point (1, 1) shows improved fusion performance. Examples are the points (2, 2), (3, 2) and (3, 3) for speaker 2 and the points (5, 3) and (7, 4).

By tuning the parameters (n, m) to any value that falls below the area of lines for the data point (1, 1), both the detection error rates can be arbitrarily reduced with a trade-off in verification time. It is shown that there is potential to improve the performance of even weaker classifiers by combining them in this manner. The FRR and FAR for speaker 9 increases up to 19.6% and 16.7% respectively by considering 4 attempts at each of the seven digit classifier combination. However, further improvement in performance is possible by increasing the number of repetitions for each sample.

Comparison of Ideal and Experimental Error Rates:

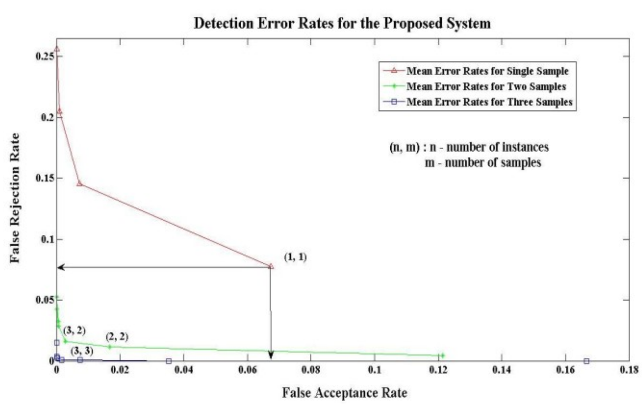
As discussed in section 2 the proposed system requires the tuning of parameters (n, m) before performing real-world verification. This tuning is usually done by estimating the desired ideal FAR and ideal FRR using the equations (5) and (6). An analysis is done here to find whether the theoretically predicted ideal FAR and ideal FRR are statistically similar to the experimentally obtained error rates.

The equations (5) and (6) are proposed assuming the error probabilities to be the same for individual classifiers. These equations can be expanded to include different error rates for each classifier.

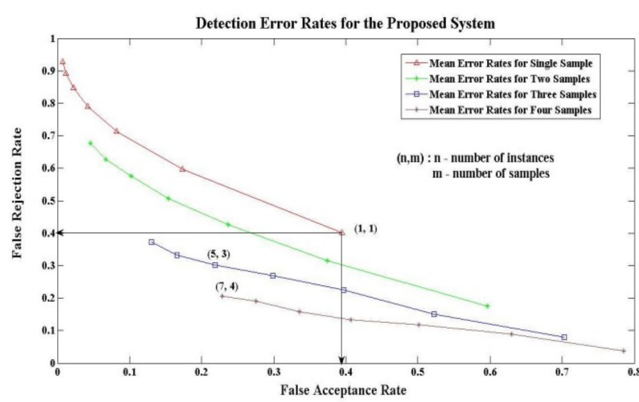
$$\alpha_{ideal} = m\alpha_1 * m\alpha_2 * ... * m\alpha_n \quad (9)$$

$$\rho_{ideal} = \rho_1^m + (1-\rho_1^m)\rho_2^m + (1-\rho_1^m)(1-\rho_2^m)\rho_3^m + .. \\ ..... + (1-\rho_1^m)(1-\rho_2^m)....(1-\rho_{n-1}^m)\rho_n^m \quad (10)$$

The ideal FAR and FRR are obtained by substituting the error rates for individual digits from the tune dataset (Table 1). It is to be noted that the experimentally obtained values here would not be exactly same as ideal values as there is a difference in classifiers performance on the tune and test data sets (Table 1).



(a)



(b)

Figure 5. Detection error rates for the proposed system - curves represent the use of multiple samples and the data points on each curve represent different classifier combinations for (a) speaker 2 (b) speaker 9

The bar graphs in the figure 6 show the comparison between the ideal and experimental detection errors for the speaker 9. The error rates for different classifier combinations with no repetition and one repetition of a sample are presented in the figure. As the individual error rates for speaker 9 are high, the ideal false acceptance rates reaches the upper bound ($\alpha_1\alpha_2...\alpha_n \leq 1/m^n$) for classifier combinations with 3 and 4 multiple samples.

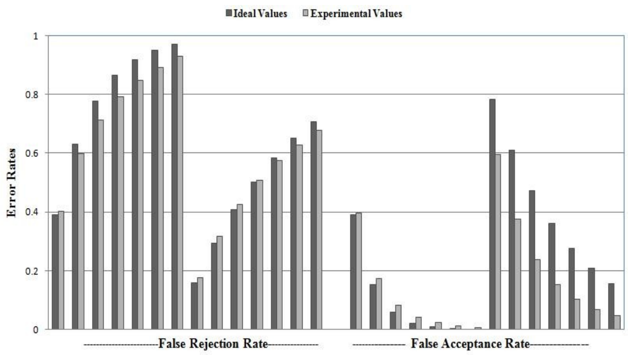


Figure 6. Comparison of Ideal and Experimental Values of FRR and FAR.

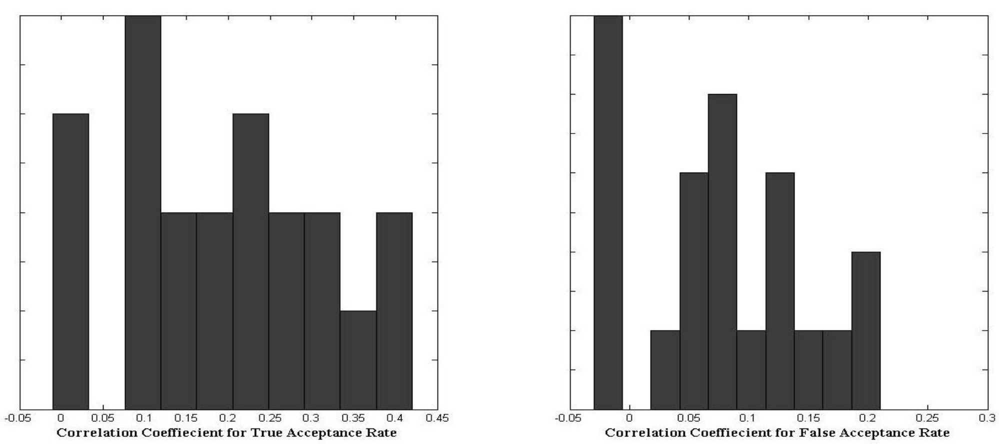
From the bar graph it is evident that there is some difference between the ideal and experimental mean detection errors. One of the predicted reasons for this difference might be that the some of the classifier decisions may be statistically dependent (and correlated) and so the error probabilities may be larger or smaller than the expressions in Equations (7) and (8) for statistically independent classifier decisions [10, 11]. The input data presented at each classifier may be correlated even though the text is different [12].

Correlation coefficients are calculated by first finding the degree of dependence between the decisions. The approach used here is based on Bahadur-Lazarsfeld expansion [11]. Figure 7(a) and 7(b) show the histogram of the correlation coefficient for true acceptance rate (TAR) and false acceptance rate (FAR) respectively for speaker 9. It is evident from the figure that the classifier decisions are correlated and further the correlation between the classifier decisions is less for impostor testing than genuine testing for speaker 9.

Kai et al. [13] explored classifier selection (selecting a subset of classifiers from a larger set) methods to achieve optimal performance using correlation analysis. It is possible to adapt this methodology for finding the optimal set of classifiers, in this case best set of digits, specific to a speaker for performance enhancement. Methods for modelling the dependencies between the classifier decisions and the classifier subset selection for optimal performance will be explored in future.

5. Conclusion and Future Work

This work demonstrates that the proposed sequential decision fusion system can be effectively used to control the detection errors. This framework of multiple sample and multiple instance combination of classifiers is analytically and experimentally evaluated using a text-dependent speaker verification system. It is shown that there is potential to improve the performance of weaker classifiers by combining them in this manner. This work also demonstrates that superior performance can be obtained despite the seemingly ideal assumption that classifiers make uncorrelated decisions. Though analysis here is done based on speech modality, the framework can be applied to handwriting, key stroke dynamics and other modal characteristics.



(a) **(b)**

Figure 7: Histograms for correlation coefficients between two classifier decisions for (a) genuine user testing (b) impostor testing

Future work will also consider the modelling of adaptation in repetitive samples and use this information for impostor detection. The role of statistical dependence and correlation between the classifier decisions will be investigated. Further, methods to obtain statistically independent information for classification will also be explored.

Acknowledgement

The authors acknowledge the contributions of Dr. Eddie Wong and D. Lawson in the QUT Vocal Access Speaker Verification software package and assistance from Prof. Sridharan and members of the SAIVT research group at QUT.

References

[1] A. Ross, *et al.*, *Handbook of multibiometrics*: Springer-Verlag New York Inc, 2006.

[2] A. Jain, *et al.*, "Fingerprint matching: Data acquisition and performance evaluation," *Dept. of Computer Science, Michigan State Univ., East Lansing, Tech. Rep. MSU-CPS-99–14,* 1999.

[3] F. Wang and J. Han, "Information fusion in personal biometric authentication based on the iris pattern," *Measurement Science and Technology,* vol. 20, p. 045501, 2009.

[4] M. Cheung, *et al.*, "Multi-sample fusion with constrained feature transformation for robust speaker verification," in *Eighth International Conference on Spoken Language Processing (ICSLP)*, Jeju Island, Korea, 2004, pp. 1813-1816.

[5] D. A. Reynolds, "An overview of automatic speaker recognition technology," in *Acoustics, Speech, and Signal Processing, 2002. Proceedings. (ICASSP '02). IEEE International Conference on*, 2002, pp. IV-4072-IV-4075 vol.4.

[6] J. de Veth and H. Bourlard, "Comparison of hidden Markov model techniques for automatic speaker verification in real-world conditions," *Speech Communication,* vol. 17, pp. 81-90, 1995.

[7] R. Kashi and W. Nelson, "Signature verification: benefits of multiple tries," in *Frontiers in Handwriting Recognition, 2002. Proceedings. Eighth International Workshop on*, 2002, pp. 424-427.

[8] V. Chandran and A. Nguyen, "Biometrics: New Perspectives on Multimodal and Client-centred Systems," in *International Workshop on Recent Advances in Biometrics*, Kanpur, India, 2005, pp. 77-89.

[9] R. Cole, *et al.*, "The CSLU speaker recognition corpus," in *Proceedings of International Conference on Spoken Language Processing*, Sydney, Australia, 1998, pp. 3167-3170.

[10] K. Venkataramani and B. V. K. V. Kumar, "Conditionally Dependent Classifier Fusion Using AND Rule for Improved Biometric Verification," in *Pattern Recognition and Image Analysis*. vol. 3687/2005, ed: Springer Berlin / Heidelberg, 2005, pp. 277-286.

[11] K. Venkataramani and B. Kumar, "Role of Statistical Dependence Between Classifier Scores in Determining the Best Decision Fusion Rule for Improved Biometric Verification," in *Multimedia Content Representation, Classification and Security*. vol. 4105/2006, ed: Springer Berlin / Heidelberg, 2006, pp. 489-496.

[12] E. N. Zois and V. Anastassopoulos, "Fusion of correlated decisions for writer verification," *Pattern Recognition,* vol. 34, pp. 47-61, 2001.

[13] K. G. a. W. Yan, "Using Correlation-Based Measures to Select Classifiers for Decision Fusion," in *in Proceedings of SPIE Defence and Security Symposium: Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications* 2005, pp. 180-191.