



Enhancement of multimodal biometric segregation using unconstrained cohort normalisation

F. Alsaade^a, A.M. Ariyaeinia^{a,*}, A.S. Malegaonkar^a, M. Pawlewski^b, S.G. Pillay^a

^aUniversity of Hertfordshire, College Lane, Hatfield, Hertfordshire, AL10 9AB, UK

^bBT Laboratories, Martlesham, Ipswich, IP5 7RE, UK

Received 30 April 2007; accepted 26 June 2007

Abstract

This paper presents an investigation into the effects, on the accuracy of multimodal biometrics, of introducing unconstrained cohort normalisation (UCN) into the score-level fusion process. Whilst score normalisation has been widely used in voice biometrics, its effectiveness in other biometrics has not been previously investigated. This study aims to explore the potential usefulness of the said score normalisation technique in face biometrics and to investigate its effectiveness for enhancing the accuracy of multimodal biometrics. The experimental investigations involve the two recognition modes of verification and open-set identification, in clean mixed-quality and degraded data conditions. Based on the experimental results, it is demonstrated that the capabilities provided by UCN can significantly improve the accuracy of fused biometrics. The paper presents the motivation for, and the potential advantages of, the proposed approach and details the experimental study. © 2007 Pattern Recognition Society. Published by Elsevier Ltd. All rights reserved.

Keywords: Multimodal biometrics; Score-level fusion; Unconstrained cohort normalisation; Biometric identification; Biometric verification

1. Introduction

The automatic recognition of the identities of individuals is becoming an increasingly important requirement in a variety of applications. Examples are teleshopping, telebanking, and physical access control. Certain obvious attractions of biometrics over the conventional means of identification have made it a superior choice in a growing range of scenarios. In recent years, an area of considerable interest in biometric recognition has been the use of multiple modalities. This is partly in view of the possibility of such limitations as non-universality and impersonation with the unimodal biometric techniques. However, a main attraction of multimodal biometrics, which is the subject of this paper, is that it provides the opportunity for enhancing the recognition accuracy beyond that achievable with unimodal biometrics.

In general, multimodal biometrics is based on the notion that the sets of data obtained from different modalities are complementary to each other [1]. Consequently, an appropriate combination of such data sets can be more useful than using the data from any single modality. For this purpose, there are various data combination levels that can be considered. Examples are the feature level, score level and decision level [1]. It has, however, been reported that the most appropriate and effective approach to multimodal biometrics is through the fusion of data at the score level [2].

In general, one of the important problems associated with any multimodal as well as unimodal technique is the undesired variations in the biometric data. Such variations are reflected in the corresponding biometric scores, and thereby can adversely influence the overall effectiveness of biometric recognition. The said variations can arise due to the effects of data capturing apparatus and various non-ideal operating conditions such as background noise and ambient lighting effects.

This paper presents investigations for enhancing the accuracy of multimodal biometrics, through the introduction of appropriate score normalisation into the field. The focus of the study is on the score-level fusion of face and voice biometrics

* Corresponding author. Tel.: +44 1707 284348; fax: +44 1707 284199.

E-mail addresses: F.Alsaade@herts.ac.uk (F. Alsaade),

A.M.Ariyaeinia@herts.ac.uk (A.M. Ariyaeinia),

A.Malegaonkar@herts.ac.uk (A.S. Malegaonkar), mark.pawlewski@bt.com

(M. Pawlewski), S.G.Pillay@herts.ac.uk (S.G. Pillay).

using SVM (support vector machine). The use of SVM in this work is based on earlier studies reporting it as one of the most effective methods for multimodal biometric fusion [1,3]. However, because of the generality of the approach proposed in this paper, the outcomes should be applicable to other fusion methods as well.

The rest of the paper is organised as follows. Section 2 introduces the proposed approach and discusses the motivation behind its use. The experimental investigations and an analysis of the results are presented in Section 3, and the overall conclusions are given in Section 4.

2. Motivation and proposed approach

An important requirement for the effective operation of a multimodal biometric system in practice is the existence of capability for minimising the effects of variations in the data from the individual modalities deployed. This would then lead to maximising the recognition accuracy in the presence of variation (e.g. due to contamination) in some or all types of biometric data involved. In reality, however, this is a challenging requirement as the data variation can be due to a variety of reasons, and can have different characteristics. Another aspect of difficulty in multimodal biometrics is the lack of information about the relative variation in the different types of biometric data.

In recent years, there has been considerable research into methods for dealing with data quality in fusion-based biometrics [4–7]. However, the work carried out to date has, in general, been concerned with adjusting the balance of weighting in fusion in favour of the modalities of better quality. In other words, emphasising or deemphasising the scores for the individual biometric modalities in the fusion process, based on an estimate of their relative degradation. The results of these studies have all verified that the introduction of an appropriate weighting scheme can be beneficial in multimodal fusion. However, it is believed that the effectiveness of multimodal biometrics can be further improved if, through some means, the scores from the degraded modalities can be corrected appropriately. According to the literature, an approach with the potential for offering the above desired capability is that of score normalisation. To date, this method has been used only in the context of speaker recognition [8,9]. The approach is based on the concept that if anomalous events in the test utterance cause a speaker's score against his (her) own model to degrade, then the scores obtained for the same speaker against certain other background models are also affected in the same way. As a result, the ratio of the score for the target model to a statistic of scores for the considered background models remains relatively unchanged. The use of this ratio instead of the absolute score for the target model has been shown to improve the verification performance.

The development of the concept of score normalisation in speaker recognition has been based on the fact that the statistical speaker classifiers provide the verification score as the probability of the observed test utterance x , given the target model λ . In other words, they compute the probability for the target model

producing the observed utterance. However, since the observed test material is in fact the test utterance, what is required to be computed is the probability of the target model, given the test utterance. These two properties are related through the Bayes' theorem as [9,10]

$$p(\lambda|x) = \frac{p(x|\lambda)p(\lambda)}{p(x)}, \quad (1)$$

where $p(\cdot)$ is the probability function. In this equation, the speaker model probability, $p(\lambda)$, can be assumed equal for all speakers, and therefore ignored. $p(x)$, on the other hand, will need to be approximated. To date, there have been different approximation approaches introduced for this purpose, leading to different score normalisation methods [9–11]. A slightly different approach to score normalisation in speaker recognition is that based on the standardisation of score distributions, with the aim to facilitate the use of a single threshold for all registered speakers [8]. A major difficulty in setting a global threshold in speaker verification (SV) is that both impostor score distribution and true speaker score distribution have different characteristics for different registered speakers. An approach to tackling this issue is that of fixing the characteristics of one of the score distribution types for all registered speakers. Currently, the common practice is to focus on standardising the impostor score distributions. The main reason for operating on the impostor score distributions, rather than on the true speaker score distributions, is the unavailability of sufficient data (in the existing databases) for a reliable estimation of the standardisation parameters in the latter approach. The different methods in this two category of score normalisation (i.e. Bayesian and standardisation) have already been subjected to thorough comparative evaluations in the context of speaker recognition [10,12]. The normalisation methods considered for this purpose are cohort normalisation (CN), unconstrained cohort normalisation (UCN), universal background model (UBM) normalisation, T-norm and Z-norm. The outcomes, which have been based on the use of decoupled reference modelling, have indicated UCN as the best performing normalisation technique. The study has also shown that whilst T-norm is amongst the best performers in speaker verification, it provides one of the worst results in the verification stage of open-set identification (OSI), even when combined with Z-norm.

The current state-of-the-art in speaker recognition, involve the use of GMM-UBM [13]. The advantage of this approach is twofold. First, it helps alleviate the adverse effects of unseen data. Second, it provides a useful means for score normalisation. However, the method requires the use of UBM-based adapted modelling which is developed specifically for speaker recognition, and is not applicable to other biometric modalities. According to the study in Ref. [14], T-norm is extremely effective for open-set speaker identification as well as speaker verification, only when speaker models are obtained by appropriately adapting a UBM. Since such adapted modelling is only feasible in the context of speaker recognition, for the purpose of consistency, both biometric modalities considered in this study are based on decoupled reference material. In this case, UCN appears as the best choice for the purpose of score

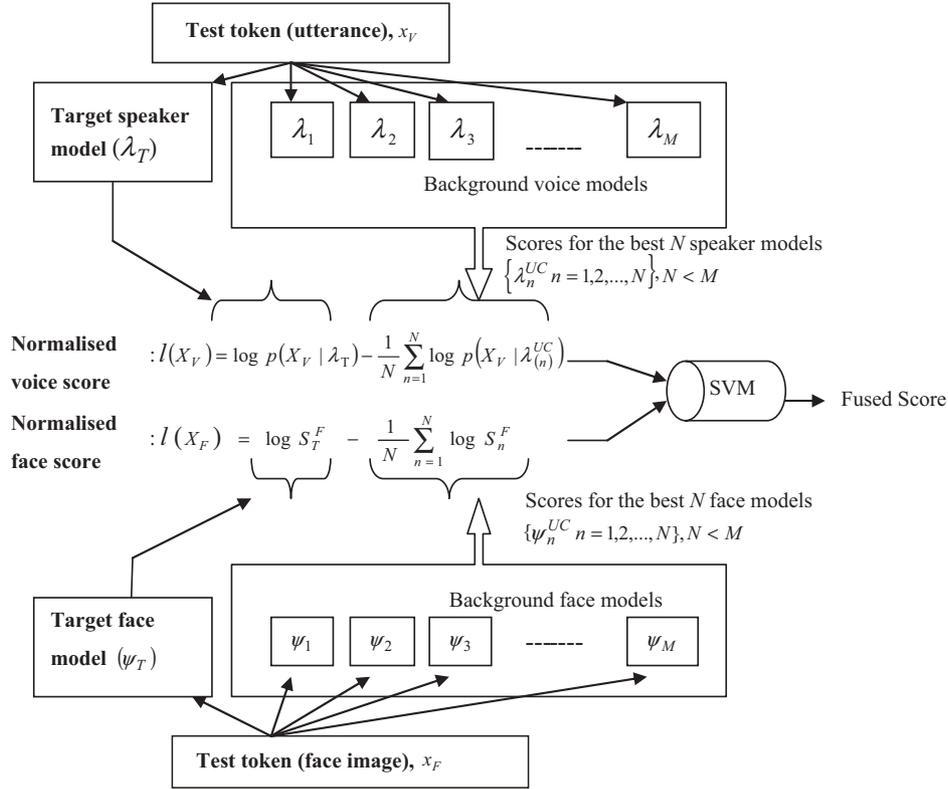


Fig. 1. Unconstrained cohort normalisation of scores in multimodal biometric fusion. Note: S_T^F and S_n^F are the scores obtained for the target model (ψ_T) and background models respectively, using the test face image.

normalisation, and therefore deployed in this study. It should be pointed out that, in general, such consistency across different modalities involved is not essential. In other words, in multimodal biometrics involving voice, the speaker representation can be based on adapted models whilst, the decoupled representation approach is used for other modalities. In such a scenario, certain other established methods may also be considered for the normalisation of speaker recognition scores, but UCN is still the most appropriate choice for modalities involving decoupled reference material.

In UCN, $p(x)$ in Eq. (1) is approximated as

$$p(x) \approx \left[\prod_{k=1}^K p(x|\lambda_k) \right]^{1/K}, \quad (2)$$

where $p(x|\lambda_k), k = 1, \dots, K$, are the top K probabilities obtained for the observation, using a set of M background speaker models ($M > K$). These top scoring models are called competing models and their selection is carried out dynamically based on their closeness to the observed utterance in the test phase.

Based on the above, the normalised score can be expressed in the log domain as

$$S_{UCN} = \log p(x|\lambda) - L(x), \quad (3)$$

where $L(\cdot) = \log p(\cdot)$. This equation suggests that the effects of data degradation can be significantly reduced if these are reflected similarly in $L(x)$ and the target model score. As already shown in Ref. [15], this approach works effectively

regardless of whether the operating framework is probabilistic or non-probabilistic. Therefore, provided UCN exhibits similar characteristics with other types of biometrics, its application to multimodal biometric fusion can be of considerable value for enhancing the reliability of the process in uncontrolled/varied operational conditions. This is because the approach provides a useful means for appropriately adjusting the individual biometric scores for a client, without any prior knowledge of the level of degradation of each biometric data type involved. However, to date, there have been no reported investigations into the use of UCN with any biometrics other than voice. The aim of this paper is therefore to explore the potential usefulness of score normalisation in an additional modality (i.e. face biometrics) and to investigate its effectiveness for enhancing the accuracy of multimodal biometrics. Fig. 1 illustrates the concept of deploying UCN in a multimodal biometric recognition scenario.

Another interesting and beneficial aspect of using UCN in multimodal biometrics is that it can potentially facilitate the separation of the scores for a given client from those for impostors targeting that client. This is based on the suppression of all the individual biometric scores for the latter in relation to those for the former. The reason is that, for a given type of biometrics and an adequately large set of background models, an impostor targeting a particular client model is likely to match one or few of the background models more closely. As a result, the application of UCN can result in reducing the impostor biometric scores relative to those of the client. The combina-

tion of the above two characteristics of UCN suggests that the technique can help enhance the biometrics reliability in both clean and adverse conditions. It is also thought that these capabilities should significantly increase the multimodal biometric accuracy. This is because the technique operates on the individual biometric scores involved independently, and the accuracy of the final fused score in multimodal recognition can benefit from the enhancement achieved in all these individual scores.

3. Experimental investigations and results

The experimental studies are concerned with the score-level fusion of face and voice biometrics in the two recognition modes of verification and OSI. The modelling and pattern matching approaches used with each modality is not discussed here, as these are outside the scope of this study. The investigations in each recognition mode involve three different data conditions. The first two are formed by using scores for clean face images together with scores for either clean or degraded utterances. The third one is based on the use of scores for degraded face images and degraded utterances.

In each experiment, the individual biometric score types involved are subjected to the range equalisation process using the Z -score normalisation [2]. In this work, the process of score-level fusion is based on the use of linear SVM [16]. The fusion process is applied to the biometric scores with and without subjecting them to the UCN process. This is to determine the level of effectiveness enhancement offered by UCN. The competing models required for UCN are selected from within the set of registered users during the test phase. The cohort size of the competing models is set to 1 and 3 in the cases of clean and degraded data, respectively. This is in agreement with the findings in some earlier studies [9,10]. The procedures for speech feature extraction and speaker classification are as detailed in Refs. [10,12]. The face recognition scores are based on the approaches detailed in Refs. [17,18].

3.1. Fusion under clean data conditions

The aim of the experiments in this part of the study is to investigate the effectiveness of UCN in enhancing the reliability of multimodal fusion when the biometric data sets are free from degradation. The data sets considered for the face and voice modalities in this investigation are extracted from the XM2VTS and TIMIT databases, respectively [17,19]. Using these biometric data sets, a total of 235 chimerical identities are formed. These consist of 140 clients, 25 development impostors and 70 test impostors. The development data comprises 140 and 22960 (i.e. $140 \times \{25 + [140 - 1]\}$) score tokens from the same-users and impostors (including cross-users), respectively. The corresponding score tokens used in the testing phase are 140 and 29260 (i.e. $140 \times \{70 + [140 - 1]\}$), respectively.

The results for the verification experiments in this part of the study are presented as equal error rates (EERs) in Table 1. As observed, the use of UCN has resulted in reducing the verification EERs for the individual modalities and for the fused biometrics. These outcomes, confirm the earlier suggestion

Table 1
Effectiveness of UCN in multimodal verification based on clean biometric data

Modality	EER% (without UCN)	EER% (with UCN)
Voice (TIMIT)	2.61	0.05
Face (XM2VTS)	3.57	2.86
Fused: voice and face	0.11	≈ 0.00

Table 2
Experimental results for open-set identification based on clean biometric data

Modality	Without UCN		With UCN	
	IER%	OSI-EER%	IER%	OSI-EER%
Voice (TIMIT)	≈ 0.00	17.14	≈ 0.00	2.86
Face (XM2VTS)	9.29	12.86	9.29	8.57
Fused: voice and face	0.71	2.86	≈ 0.00	≈ 0.00

(Section 2) that the use of UCN in clean data conditions is still beneficial. The effectiveness of UCN under such operating condition is due to its ability to suppress the scores for impostors in relation to those for true users. It is noted that the usefulness of UCN in fused biometrics is mostly due to its performance with the voice modality. However, the corrective effect that UCN has on the face modality is seen to be also considerable. This in turn has helped further enhance the accuracy of classification based on the fused data. It should be emphasised that this is the first time that the use of UCN with face biometrics has been investigated and its effectiveness demonstrated.

Table 2 presents the results of OSI experiments with clean data. These are expressed in terms of IER (identification error rate) and OSI-EER that occur in the first and second stages of the process, respectively. An interesting aspect of these results is that the use of UCN does not change the IER for any of the single modalities whilst, it successfully reduces IER (to zero in this case) for the fused biometrics. The reason for this phenomenon can be described as follows. Firstly, like any other score normalisation method, UCN cannot be expected to correct any misidentification occurring in the first stage of unimodal OSI [12]. However, what is achieved through UCN is the suppression of the scores which lead to the misidentification in the individual modalities, in relation to the scores for the correct identities. Although this does not lead to the re-ranking of the unimodal identity scores, it facilitates the reduction of misidentification in the fusion stage. It is also interesting to note that, in this case, the use of UCN appears to ensure that the lowest error rates are obtained through the fused biometrics.

The results in Tables 1 and 2 indicate that the OSI-EERs in unimodal biometrics are considerably larger than the EERs for the verification experiments. This is due to the fact that the verification stage in OSI is more challenging than the standard biometric verification [10]. The reason is that in the former process, each unknown (unregistered) user will need to be discriminated from his/her best matched registered user. In other words, the verification stage in OSI can be considered as a specific (but unlikely) scenario in the standard verification process

Table 3
Performance of UCN in biometric verification based on mixed-quality data

Modality	EER% (without UCN)	EER% (with UCN)
Voice (NIST)	26.24	10.00
Face (XM2VTS)	3.57	2.86
Fused: voice and face	2.86	0.78

Table 4
Experimental results for open-set identification based on mixed-quality data

Modality	Without UCN		With UCN	
	IER%	OSI–EER%	IER%	OSI–EER%
Voice (NIST)	40	45.71	40	15.71
Face (XM2VTS)	9.29	12.86	9.29	8.57
Fused: voice and face	6.43	12.86	4.29	5.71

in which each impostor targets only his or her closest model in the registered set.

In multimodal biometrics, however, it is very unlikely that different biometric modalities of an impostor are best matched to the corresponding modalities of an individual registered user. Consequently (as the experimental results show), fusing the biometric scores leads to a significant improvement in the verification accuracy. The use of UCN in this case is observed to maximise the fused biometrics accuracy as well as considerably reducing the OSI–EER for each of the modalities involved. As indicated earlier, this is achieved through UCN suppressing the scores for unknown users in relation to those of registered users.

3.2. Fusion under varied data quality conditions

The purpose of the experiments presented in this section is to investigate the usefulness of UCN in multimodal fusion when the qualities of the biometric data types are considerably different. The data sets considered for the face and voice modalities in this case are extracted from the XM2VTS (clean images) [17] and from the 1-speaker detection task of the NIST Speaker Recognition Evaluation 2003 (degraded speech) databases, respectively [12]. Using these data sets, again a total of 235 chimerical identities are formed. These consist of the same number of clients, development impostors and test impostors as in the previous experiments (Section 3.1). The development and test data sets also consist of the same number of score tokens from the same-users and impostors as those considered in the previous section.

The results of verification and OSI in this case are presented in Tables 3 and 4, respectively. It is noted that whilst the error rates for the face modality are exactly the same as those in the previous investigation, due to the use of a degraded speech database, the accuracy rates for the voice modality in this case are lower than the corresponding ones in Section 3.1. The results in Table 3 demonstrate the capability of UCN in reducing the verification error rate, particularly, in fused biometrics. UCN achieves this by a combination of enhancing the client scores

Table 5
Effectiveness of UCN in multimodal verification based on degraded data

Modality	EER% (without UCN)	EER% (with UCN)
Voice (NIST)	35.69	15.38
Face (BANCA)	18.27	13.46
Fused: voice and face	16.35	6.35

when these are affected by data degradation, and suppressing the impostor scores in relation to the client ones. It is noted that without UCN, the fusion process results in improving the EER associated with the better modality by about 20%. According to the results, this reduced EER (2.86%) is further decreased by about 73% through the use of UCN.

It is observed from the results in Table 4 that the use of fusion process, in this case, leads to reducing the lowest IER offered by unimodal biometrics. However, it is also seen that this capability of fused biometrics is considerably improved through UCN. On the other hand, it is observed that, in this case, the fusion process can reduce the OSI–EER% only when used together with UCN. The reduction in OSI–EER achieved with such a combination is in excess of 55%.

Another important outcome of the experimental investigations can be observed by considering the results in Table 4 together with those in Table 2. Based on these results, it is clearly seen that the fusion process on its own may not necessarily lead to the reduction of IER or OSI–EER offered by the best single biometric modality involved. The results in these two tables indicate that it is through the deployment of UCN that the fused biometrics can consistently outperform unimodal biometrics.

3.3. Fusion under degraded data conditions

The experiments in this section investigate the effectiveness of UCN in enhancing the reliability of multimodal fusion when the two biometric data types adopted are both degraded. The data set for the face modality in this investigation is extracted from the BANCA (degraded images) database [18] whilst the data for the speech modality is extracted from the 1-speaker detection task of the NIST Speaker Recognition Evaluation 2003 (degraded speech) database [12]. Using these biometric data sets, a total of 52 chimerical identities consisting of 26 clients and 26 impostors are formed. The face recognition scores are obtained based on images captured in four sessions, and affected by two different forms of distortions [18]. Based on these and the corresponding score data for NIST, a development score data set is formed for the experiments. This consists of 104 and 5304 score tokens from the same-users and impostors (including cross-users), respectively. The corresponding score tokens used in the testing phase are also 104 and 5304, respectively. Tables 5 and 6 present the results obtained in this case for verification and OSI, respectively.

It can be seen from the experimental results in Table 5 that the use of UCN has again resulted in the reduction of the verification EERs for the individual modalities as well as for the fused biometrics. It can also be observed that the fusion process

Table 6
Experimental results for open-set identification based on degraded biometric data

Modality	Without UCN		With UCN	
	IER%	OSI-EER%	IER%	OSI-EER%
Voice (NIST)	26.92	48.08	26.92	15.38
Face (BANCA)	38.46	30.77	38.46	25.00
Fused: voice and face	25.00	31.73	18.27	9.62

on its own outperforms the best individual modality involved. On the other hand, it is seen that the verification accuracy offered by fused biometrics is increased significantly (by about 61%) through the use of UCN prior to fusion. It is worth noting that the accuracy of fused biometrics without UCN (Table 5) is below the accuracy obtained by using UCN with any of the two single modalities involved. These results are in agreement with the earlier suggestions (Section 2) that the use of UCN in degraded data conditions is beneficial. The effectiveness of UCN under such operating conditions is due to the twofold characteristic of UCN. Firstly, it provides a means for enhancing the scores when the test data is degraded, and secondly, it aims to suppress the scores from impostors in relation to those for clients.

In Table 6, it is observed that the fusion process results in an IER which is slightly better than the IER offered by the best unimodal biometrics. However, using UCN together with the fusion process leads to a considerably lower IER. It is also observed that, in this scenario, the fusion process reduces the OSI-EER only when used in conjunction with UCN. In fact, without UCN, the OSI-EER obtained with fused biometrics is worse than that for the better of the two modalities. The use of UCN is seen to reduce the OSI-EER for the fused biometrics by about 70%. Again it is noted that, in terms of OSI-EER, the performance of fused biometrics without UCN is well below that of either of the modalities with UCN. In brief, the results in this study indicate that it is only through the deployment of an appropriate score normalisation technique, in this case UCN, that the fused biometrics can consistently outperform the unimodal biometrics involved.

Figs. 2 and 3 further illustrate the results obtained for the verification and the second stage of OSI experiments in this part of the study, respectively. Fig. 2 clearly shows the significant increase in the reliability of fused biometrics obtained through the use of UCN. The plots in this figure also illustrate the considerable performance improvements achieved through the use of UCN with the individual modalities, which is the cause of the above mentioned enhancement in the accuracy of fused biometrics.

The DET plots in Fig. 3 further emphasise the role of UCN in enhancing the reliability of fused biometrics. In fact, it is observed that, without UCN, the fused biometrics accuracy is highly influenced by the worse of the two modalities involved and does not even match the performance of the better of the two individual modalities. On the other hand, by applying UCN to the individual modalities, the fusion process is observed to provide the highest reliability in the experiments.

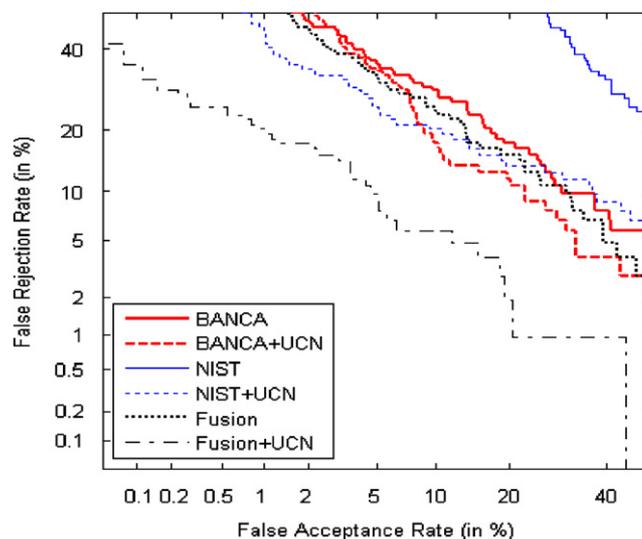


Fig. 2. DET plots for the verification experiments with degraded data.

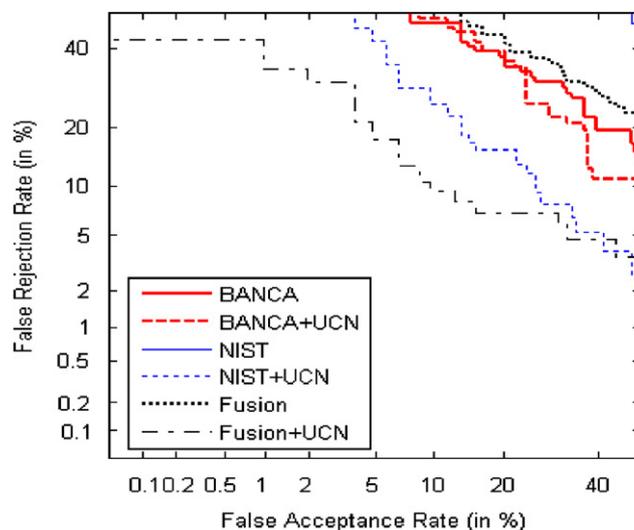


Fig. 3. DET plots for the verification process in the second stage of open-set identification experiments with degraded data. *Note:* the plot for NIST data set (without UCN) is mostly outside the scale due to the excessive high error rate in this case.

4. Conclusion

This paper has presented an investigation into the use of unconstrained cohort normalisation (UCN) with score-level fusion for multimodal biometrics. Based on the experimental investigations, it has been shown that UCN offers considerable improvements to the accuracy of multimodal biometrics in both degraded and clean data conditions. This is shown to be due to the twofold characteristic of this score normalisation method. Firstly, it provides a means for enhancing the scores when the test data is degraded, and secondly, it aims to suppress the scores from impostors in relation to those for clients. The investigations have also confirmed the usefulness of UCN in face recognition as well as in speaker recognition for which the

technique had originally been developed. Additionally, through a set of open-set identification experiments, it has been shown that multimodal fusion can consistently outperform the accuracy offered by the best single modality performer, when it is combined with UCN. Whilst this study has confirmed the effectiveness of UCN for face modality as well as fused voice and face biometrics, further investigations are required to determine what other types of biometrics can benefit from such forms of score normalisation.

Acknowledgement

The authors wish to express their gratitude to the Aristotle University of Thessaloniki and the University of Vigo and, in particular, Professor Ioannis Pitas, Professor Carmen Garcia Mateo, Professor Jose Luis Alba, Mr. Stefanos Zafeiriou, and Mr. Daniel Gonzalez for their support and provision of the face recognition scores.

References

- [1] A.K. Jain, A. Ross, S. Prabhakar, An introduction to biometric recognition, *IEEE Trans. Circuits Syst. Video Technol.* 14 (2004) 4–19.
- [2] M. Indovina, U. Uludag, R. Snelick, A. Mink, A. Jain, Multimodal biometric authentication methods: a COTS approach, in: *Proceedings of the MMUA, Workshop on Multimodal User Authentication*, Santa Barbara, CA, December 11–12, 2003, pp. 99–106.
- [3] B. Gutschoven, P. Verlinde, Multi-modal identity verification using support vector machines (SVM), in: *Proceedings of the International Conference on Information Fusion, FUSION*, IEEE Press, New York, 2000, pp. 3–8.
- [4] J. Fierrez-Aguilar, J. Ortega-Garcia, J. Gonzalez-Rodriguez, J. Bigun, Discriminative multimodal biometric authentication based on quality measures, *Pattern Recognition* 38 (2005) 777–779.
- [5] F. Alsaade, A. Ariyaeinia, L. Meng, A. Malegaonkar, Multimodal authentication using qualitative support vector machines, *INTERSPEECH* 2006, September 17–21 (2006) 2454–2457.
- [6] J. Bigun, J. Fierrez-Aguilar, J. Ortega-Garcia, J. Gonzalez-Rodriguez, Multimodal biometric authentication using quality signals in mobile communications, in: *Proceedings of the IAPR International Conference on Image Analysis and Processing, ICIAP'03*, Mantova, Italy, September 2003.
- [7] N. Poh, S. Bengio, Improving fusion with margin-derived confidence in biometric authentication tasks, in: *Fifth AVBPA*, Rye Brook, July 2005, pp. 474–483.
- [8] R. Auckenthaler, M. Carey, H. Lloyd-Thomas, Score normalization for text-independent speaker verification systems, *Digital Signal Process.* 10 (2000) 42–54.
- [9] A. Ariyaeinia, P. Sivakumaran, Analysis and comparison of score normalisation methods for text-dependent speaker verification, in: *Proceedings of the Eurospeech'97*, vol. 3, September 1997, pp. 1379–1382.
- [10] A. Ariyaeinia, J. Fortuna, P. Sivakumaran, A. Malegaonkar, Verification effectiveness in open-set speaker identification, in: *IEE Proc. Vision Image Signal Process.* 153 (5) (2006) 618–624.
- [11] D. Reynolds, Comparison of background normalisation methods for text-independent speaker verification, in: *Proceedings of the Eurospeech 1997*, Rhodes, 1997, pp. 963–966.
- [12] J. Fortuna, P. Sivakumaran, A. Ariyaeinia, A. Malegaonkar, Relative effectiveness of score normalization methods in open-set speaker identification, in: *Proceedings of the Speaker Odyssey*, 2004, pp. 369–376.
- [13] D. Reynolds, T. Quatieri, R. Dunn, Speaker verification using adapted Gaussian mixture models, *Digital Signal Process.* 10 (2000) 19–41.
- [14] J. Fortuna, P. Sivakumaran, A. Ariyaeinia, A. Malegaonkar, Open-set speaker identification using adapted Gaussian mixture models, *INTERSPEECH* 2005, September 4–8 (2005) 1997–2000.
- [15] A.M. Ariyaeinia, P. Sivakumaran, M. Pawlewski, M.J. Loomes, Dynamic weighting of the distortion sequence in text-dependent speaker verification, in: *Eurospeech'99*, pp. 967–970.
- [16] C.J.C. Burges, tutorial on support vector machines for pattern recognition, *Data Min. Knowl. Discovery* 2 (1998) 955–974.
- [17] S. Zafeiriou, A. Tefas, I. Buciu, I. Pitas, Exploiting discriminant information in non-negative matrix factorization with application to frontal face verification, Presented at *IEEE Transactions on Neural Networks*, May 2006.
- [18] S. Bengio, F. Bimbot, J. Mariethoz, V. Popovici, F. Poree, E. Bailly-Bailliere, G. Mate, B. Ruiz, Experimental protocol on the BANCA database, Technical Report IDIAP-RR 02-05, IDIAP, 2002.
- [19] F. Alsaade, A. Malegaonkar, A. Ariyaeinia, Fusion of cross stream information in speaker verification, in: *Proceedings of the COST 275 Workshop on Biometrics on the Internet*, Hatfield, UK, October 2005, pp. 63–66.

About the Author—FAWAZ ALSAADE received the degrees of BSc and MSc in computer science from King Abdulaziz University, Saudi Arabia and Western Michigan University, Kalamazoo, Michigan, USA in 1998 and 2002, respectively. Over the last four years, he has been engaged in research in the field of multimodal biometrics for the degree of PhD.

About the author—ALADDIN M. ARIYAEINIA received the degrees of BEng in Communication Engineering, MSc in Digital Signal Processing, and PhD in Active Imaging in 1976, 1982 and 1986, respectively. Ariyaeinia is currently a Reader in the Faculty of Engineering and Information Sciences at the University of Hertfordshire, where he leads the Audio Processing and Biometrics Group.

About the author—AMIT S. MALEGAONKAR received the degrees of BEng (Electronics and Telecommunications) from University of Pune, India and subsequently MSc (Data Communications and Networks) and PhD in Voice Biometrics from University of Hertfordshire, in 2000, 2002 and 2006, respectively. He is currently working as a Sr. software developer in the area of speech and audio coding.

About the author—MARK PAWLEWSKI is a Technical Group Leader in the BT Security Research Centre. He has several years' experience in core algorithm development for speech recognition and speaker verification. Pawlewski is a member of the International Biometrics Standards Committee (SC37) and is currently working on a speaker recognition interchange format.

About the author—SUROSH G. PILLAY received the first class BEng (Hon) degree in Digital Communication and Electronic Engineering from the University of Hertfordshire in 2006. He is currently engaged in a research programme in the field of speaker recognition and multimodal biometrics for the degree of PhD.