PROBABILISTIC FUSION OF REGIONAL SCORES IN 3D FACE RECOGNITION

Nesli Erdogmus, Lionel Daniel, Jean-Luc Dugelay

Multimedia Communications Department, EURECOM
Sophia-Antipolis, France
{nesli.erdogmus, lionel.daniel, jean-luc.dugelay}@eurecom.fr

ABSTRACT

Information fusion in biometrics mostly relates to multi-biometric systems which attempt to improve the performance of individual matchers: multi-sensor, multi-algorithm, multimodal, etc. However, in addition to these scenarios, the need for methods to fuse the individual regional classifiers has also emerged, due to the increasing number of region-based methods proposed to overcome expression and occlusion problems in face recognition. In this paper, we present a combination approach by converting the regional match scores into probabilities with the help of estimated regional confidence measures. Initially, face is broken into several segments and similarity and confidence scores are obtained. Then, the posteriori probabilities of the user being genuine are calculated in each region given these two scores. For this calculation, the conditional densities are obtained on the training samples by applying non-parametric kernel density estimation separately for different intervals of confidence levels. Experimental results demonstrate that the inclusion of the regional confidence measures via probabilistic conversion is much more advantageous when compared to weighted sum of original scores.

Index Terms— Region-based, score fusion, face recognition, probabilistic approach

1. INTRODUCTION

Automatic identification and verification of humans has become a substantial research field today. As the need for security applications grows continuously, biometrics attracts rising attention with its reliable and efficient identity management. However, the performance of a biometric system employing a single trait is constrained by several factors: feature overlaps, noisy inputs, sensor failures, etc. [1] In order to overcome these issues, the use of multiple biometric sources was proposed in [2] and has been thoroughly investigated since then, especially on the domain of score normalization and fusion [3,4].

In this paper, we study fusion for region-based face recognition methods which prove useful in the case of intra-class variations [5-7]. Among numerous biometric traits, face stands out with its favorable reconcilement between accessibility and reliability. Face recognition is applicable at relatively high distances and without user cooperation. However, despite the decades of dedicated research, face recognition still suffers from intra-class variation problems due to various factors in real-world scenarios such as illumination, pose, expression and occlusion.

Region-based approaches suggest dividing the faces into multiple parts for more robust comparison. For instance, in [8], 2D facial image is divided into a number of sub-regions and AdaBoost algorithm is utilized to generate a strong classifier from the combination of orthogonal component principal analysis features.

Analysis of face in multiple regions has been exploited more intensely in 3D modality. 3D face recognition offers superiority over its 2D counterpart by being intrinsically robust against illumination and pose variations. However, presence of facial expressions and occlusion still deteriorate its performance. Due to being local distortions on the facial mesh, these variations are often proposed to be handled via region-based methods. In [9], facial surface region is segmented in regions according to their degree of deformation and elasticity and then, the computed similarity scores are fused using a weighted similarity metric which attaches more importance to more static regions. Spreeuwers [10] and Faltemier et al. [6] adopt similar approaches where around 30 overlapping regions on the facial surface are independently matched and the performances of numerous fusion techniques such as majority voting, Borda count, sum, product and min rules are compared. Another set of studies base their region selection on the assumption that upper parts of the face, especially around nose, are less affected by expression variations. In [11] diagonal profiles of nose region are compared for recognition. In [12] regional dissimilarity around the nose is calculated by taking the root mean square distance between point pairs after the point correspondence is achieved via Iterative Closest Point (ICP) [13] algorithm. Additionally, in [14] 3 regions, namely forehead, nose and left eye, are utilized for Principal Component Analysis and similarity metrics are combined using sum rule and Linear Discriminant Analysis. Similarly, in [15], eyes, forehead and nose regions are matched for recognition.

In this paper, prior to fusion, we propose to obtain posteriori probabilities, P(genuine|s), of being genuine given the regional matching scores (s). However, as expressed in [4], the outputs of individual classifiers are better combined directly without being converted into probabilities in the absence of confidence measures (c) which assess the nature of the input samples. By taking this into consideration, we obtain P(genuine|s) by computing the conditional densities P(s|genuine) and P(s|impostor) based on the regional confidence scores estimated via our method, previously proposed in [16].

The rest of this paper is structured as follows: In the next section the theoretical framework for probability fusion at match score level is described in detail. In Section 3, we briefly present our regional confidence estimation method and in Section 4, confidence-measure based estimation of conditional probability densities is explained. Experimental results are given in Section 5 and finally, the paper is concluded in Section 6.

2. PROBABILITY FUSION AT MATCH SCORE LEVEL

We can represent the fusion of regional match scores as the classifier combination problem formulated in [17]: A face is to be assigned to one of the *m* possible classes in the gallery based on distinct measurements (match scores - s_i) obtained from its *R* sub-regions. In our approach, this is simplified by converting the class assignment to a binary decision which answers if the match score is "*genuine*", i.e. it is computed between two samples of the same subject or "*impostor*". These two classes are modeled by the probability density functions $p(s_i|\text{genuine})$ and $p(s_i|\text{impostor})$ and their a priori probabilities of occurrence are assumed to be equal. Under this assumption, using the Bayes theorem, $P(\text{genuine}|s_1,s_2,...,s_R)$ is expressed as:

$$P(\text{genuine}|s_1,s_2,\ldots,s_R)=\frac{p(s_1,s_2,\ldots,s_R|\text{genuine})}{p(s_1,s_2,\ldots,s_R)} \quad (1)$$

where

$$p(s_1,s_2,..,s_R) = p(s_1,s_2,..,s_R|\text{genuine}) + p(s_1,s_2,..,s_R|\text{impostor}) \quad (2)$$

As suggested in the same study by Kittler et al., a posteriori probability of being genuine given R measurements is expressed in terms of decision support computations and possible dependence on joint probability density functions is ignored. This assumption of conditional statistical independence gives way to combining the a posteriori probabilities obtained from different regions by means of a product rule. However, sum approximation of the product rule, despite its foundation on highly unrealistic assumptions, is proven to be more resilient to estimation errors and hence, is adopted in this study (Equation 3).

$$P(\text{genuine}|s_1,s_2,\ldots,s_R) \approx \sum_{i=1}^{R}\frac{p(s_i|\text{genuine})}{p(s_i|\text{genuine})+p(s_i|\text{impostor})} \quad (3)$$

3. REGIONAL CONFIDENCE ESTIMATION

In this section, we will briefly explain our regional confidence estimation method for 3D facial surfaces presented in [16]. In our method, for each segment of the face, the vertices are labeled as one of the 12 primitive categories [18] and the distribution of these labels are taken as shape descriptors to estimate the confidence levels. This is achieved by an Artificial Neural Network (ANN) which is trained using automatic measurements of regional quality scores (Here, quality implies the absence of occlusions or expressions): For every face model in the training set, each region is registered to its neutral and clean equivalent belonging to the same person, via the ICP algorithm [13]. ICP registration errors between the "high quality" samples and the regions under analysis are accepted as metrics of imperfection. The ANN maps between the extracted shape descriptors and the quality measurements.

Firstly, the primitive shape class for each vertex on the facial surface is determined as one of the following: peak, pit, ridge, ravine, ridge saddle, ravine saddle, convex hill, concave hill, convex saddle hill, concave saddle hill, slope hill, flat. Then, histogram distributions of these shapes for each region are calculated, resulting in shape descriptors of size [12x1].

Subsequently, each region of all facial scans of a subject in the training set is registered to the corresponding region of a neutral (without expression) and clean (without occlusion) scan of the same person (reference model) using ICP. Inverse of the final registration errors obtained are accepted as the regional quality measurements (Figure 1).

Lastly, an ANN is trained for each region to estimate the confidence score from the given primitive shape distributions.

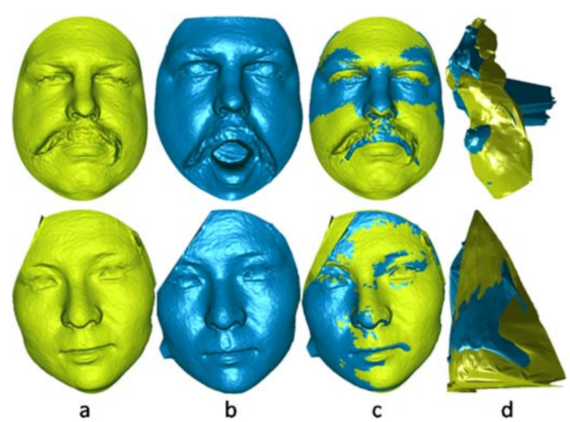

Figure 1: Examples for a bad and a good quality region: (a) neutral and clean reference model; (b) models to be evaluated; (c) model pairs after initial registration; (d) close-up to mouth to region for the 1st example and to forehead region for the 2nd example. The scores computed are 0.13 and 1.04, respectively.

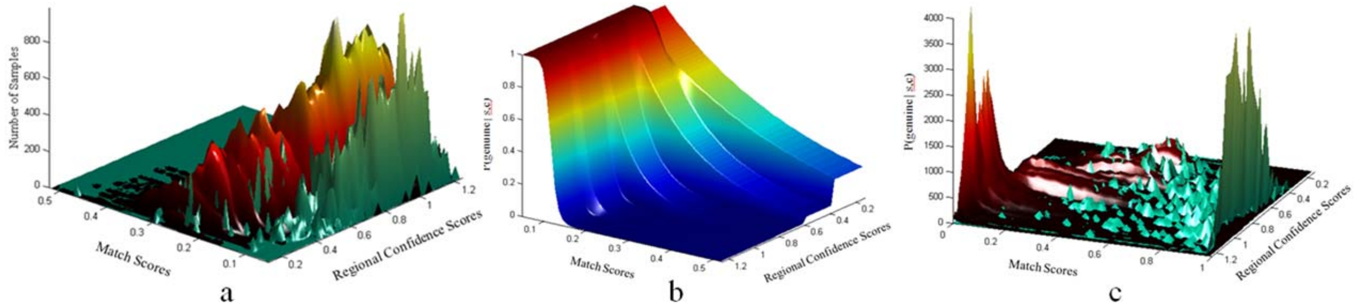

Figure 2: For the mouth region: (a) Histogram of raw match scores with respect to the regional confidence scores for mouth region: impostor scores in red and genuine scores in green (b) Calculated a posteriori probabilities for genuineness for 5 confidence level bins (c) the match score distributions after converting them to probabilities

4. CONDITIONAL DENSITY ESTIMATION

As explained in Section 2 and formulated in Equation 3, conditional probability density functions, p(s|genuine) and p(s|impostor) are essential to compute the posterior probability of genuineness.

In [19], Snelick et al. utilize mean and variance of genuine and impostor scores in the training set and assume a normal distribution for the conditional densities. However, this assumption may not be true in many cases. For this reason, Jain et al. [4] propose the use of the Parzen window based non-parametric density estimation method and obtain actual conditional densities of genuine and impostor scores.

In our work, the Parzen window is replaced with a smooth Gaussian kernel function in order to avoid discontinuities in the estimations and to take the sample point distances to the estimation point into consideration. The bandwidth is empirically set to 0.03.

Additionally, considering that as the match score increases the probability of being genuine should also increase, a monotonic sigmoidal function (Equation 4) is fit on the kernel density estimation via nonlinear regression.

$$f(x)=\frac{1}{1+e^{a*(b-x)^{c}}} \tag{4}$$

The conditional probability densities are estimated with and without considering regional confidence scores for comparison reasons. In case, where the estimated confidence levels (c) are taken into account, P(genuine|s,c) is approximated by dividing confidence scores in several bins and estimating p(s|genuine) and p(s|impostor) for each bin separately. Impostor and genuine score distributions differ greatly for different levels of region quality. With our approach, better estimates are achieved with discontinuous density functions in c-axis.

5. EXPERIMENTS AND RESULTS

We evaluated the proposed algorithm on FRGC v2 database [20]. Since a neutral and clean 3D sample is required for each person to measure the regional surface qualities automatically, the subjects without a reference model are eliminated. This results in a set of 3123 face models of 343 subjects. The first half of this dataset is utilized for two purposes: ANN training for confidence score estimations and conditional density estimations.

Face is analyzed in 7 segments: forehead, left eye, right eye, left cheek, right cheek, mouth and nose. Firstly, match scores are calculated between all training samples by employing warping parameters extracted from a TPS-based algorithm [21] as biometric signatures. Next, regional qualities are measured automatically by ICP-based registration of each facial component to the corresponding neutral and clean reference model and the primitive shape histograms are computed to train ANNs.

For each region, conditional distributions are estimated after the confidence scores are binned into several equally spaced containers. The conditional probability densities are estimated for different number of bins for comparison: 1(no bins), 2, 3, 4 and 5. Taking mouth as the exemplar region, the distribution of match scores according to the regional confidence scores, the estimated posterior probabilities P(genuine|s,c) using 5 bins and the match score distributions after being converted to probabilities are given in Figure 2.

Verification and identification performances are obtained for the proposed fusion scheme with 1, 2, 3, 4 and 5 bins using the sum rule. Success rates of the original match scores are calculated by the same fusion technique. Additionally, in order to provide equal advantage to the original scores, the regional confidence levels are incorporated via weighted sum combination:

$$S=\frac{\sum_{r=1}^{7} s_r * c_r}{\sum_{r=1}^{7} c_r} \tag{5}$$

Table 1. Comparative results for verification and identification tests before and after probability conversion of match scores

	method	VR	EER	IR
original	sum	63.39%	0.109	90.33%
original	w. sum	68.22%	0.088	91.04%
probability	1 bin	66.83%	0.072	89.76%
probability	2 bins	71.38%	0.049	91.61%
probability	3 bins	72.66%	0.047	92.13%
probability	4 bins	74.08%	0.046	91.93%
probability	5 bins	73.15%	0.045	92.00%

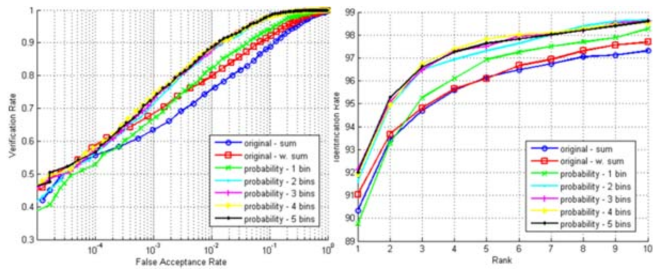

Figure 3: ROC and CMC curves for different fusion methods

The verification rates at 0.001 FAR, the equal error rates and the rank-1 identification rates for all experiments are given in Table 1. Moreover, in Figure 3, the verification and identification performances of the fusion techniques are evaluated with receiver operation characteristics (ROC) and cumulative match characteristics (CMC), respectively.

The results show that utilization of confidence scores for regional classifiers is more advantageous with the probabilistic fusion approach than employing the raw scores directly. The performances improve as the number of the bins increase but tend to converge after 3 bins. This is mainly due to scarcity of samples and hence erroneous conditional density estimations as the bins get smaller.

6. CONCLUSION

In this paper, we proposed to incorporate regional confidence scores to the fusion process via the probabilistic framework developed by Kittler et al. [17] through confidence-score based conditional density estimations. Extensive experiments conducted on the FRGC v2 database reveal that this approach can make better use of the estimated confidence levels compared to utilization of raw scores with weighted sum method.

For future work, we plan to work on non-parametric estimation of actual joint multivariate densities and utilize them for probability conversions instead of their bin-based approximations.

7. REFERENCES

[1] A. A. Ross, K. Nandakumar, and A. K. Jain, *Handbook of Multibiometrics*, Springer, 2006.

[2] M. Most, “Battle of the Biometrics”, *Digital ID World Magazine,* pp. 16-18.

[3] P. Verlinde, P. Druyts, G. Cholet, M. Acheroy, “Applying Bayes Based Classifiers For Decision Fusion In A Multi-Modal Identity Verification System,” *Proceedings of International Symposium on Pattern Recognition* “In Memoriam Pierre Devijver”, Brussels, Belgium, 1999.

[4] A. Jain, K. Nandakumar, and A. Ross, "Score Normalization in Multimodal Biometric Systems," *Pattern Recognition*, vol. 38, no. 12, pp. 2270-2285, Dec. 2005.

[5] S. H. Lin, S. Y. Kung, and L. Lin, "Face recognition/detection by probabilistic decision-based neural network," *IEEE Trans. on Neural Networks*, vol. 8, no. 1, pp. 114-132, 1997

[6] T. C. Faltemier, K. W. Bowyer, P. J. Flynn, "A Region Ensemble for 3-D Face Recognition," *IEEE Trans. on Information Forensics and Security*, vol.3, no.1, pp.62-73, 2008

[7] N. Alyuz, B. Gokberk, L. Akarun, "Regional Registration for Expression Resistant 3-D Face Recognition," *IEEE Trans. on Information Forensics and Security*, vol.5, no.3, pp.425-440, 2010

[8] J. Li, G. Poulton, Y. Guo, R.-Y. Qiao, “Face Recognition Based on Multiple Region Features,” *Proc. VIIth Digital Image Computing: Techniques and Application,* pp. 69-78, 2003

[9] B. Ben Amor, M. Ardabilian, L. Chen, "Toward A Region-Based 3D Face Recognition Approach," *IEEE International Conf. on Multimedia and Expo,* pp.101-104, 2008

[10] L. Spreeuwers, "Fast and accurate 3D face recognition," *International Journal of Computer Vision*, pp. 1-26, Mar. 2011.

[11] B. Bornak, S. Rafiei, A. Sarikhani, A. Babaei, "3D Face Recognition By Used Region-Based With Facial Expression Variation," *International Conf. on Signal Processing Systems*, vol.1, pp.V1-710-V1-713, 5-7 July 2010

[12] D. Smeets, T. Fabry, J. Hermans, D. Vandermeulen, P. Suetens, "Fusion of an Isometric Deformation Modeling Approach Using Spectral Decomposition and a Region-Based Approach Using ICP for Expression-Invariant 3D Face Recognition," *International Conf. on Pattern Recognition*, pp. 1172-1175, 23-26 Aug. 2010

[13] P. J. Besl and H. D. Mckay, "A Method For Registration Of 3-D Shapes," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* vol. 14, no. 2, pp. 239-256, 1992.

[14] Wei-Yang Lin, Ming-Yang Chen, K. R. Widder, Yu Hen Hu, N. Boston, "Fusion of Multiple Facial Regions for Expression-Invariant Face Recognition," *IEEE Workshop on Multimedia Signal Processing*, pp. 426-429, 1-3 Oct. 2007

[15] A. Mian, M. Bennamoun and R. Owens, "Region-based Matching for Robust 3D Face Recognition", *British Machine Vision Conf.*, vol. 1, pp. 199--208, 2005.

[16] N. Erdogmus, J-L. Dugelay, “Regional Confidence Score Assessment for 3d Face,” *IEEE International Conf. on Acoustics, Speech and Signal Processing,* Kyoto, 2012.

[17] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, "On combining classifiers," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* vol. 20, no. 3, pp. 226-239, Mar. 1998.

[18] O. D. Trier, T. Taxt, and A. K. Jain. “Data Capture From Maps Based On Gray Scale Topographic Analysis”, *International Conf. on Document Analysis and Recognition*, 923-926, 1995.

[19] R. D. Snelick, M. D. Indovina, J. H. Yen, and A. Mink, “Multimodal Biometrics: Issues in Design and Testing,” *ACM International Conf. on Multimodal Interface,* 2003

[20] P. J. Phillips, P. J. Flynn, T. Scruggs, K.W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, “Overview Of The Face Recognition Grand Challenge,” *IEEE Conf. on Computer Vision and Pattern Recognition*, 947-954, 2005.

[21] N. Erdogmus, J-L. Dugelay, “On discriminative properties of warping parameters for 3d face recognition,” *International Conf. on Informatics, Electronics & Vision,* Bangladesh, 2012.