
# ROBUST MULTI-MODAL AND MULTI-UNIT FEATURE LEVEL FUSION OF FACE AND IRIS BIOMETRICS

*Ajita Rattani and Massimo Tistarelli*

December 16, 2008

University of Sassari

Computer Vision Laboratory

Technical Report No. CVL -2008-003

UNIVERSITY
of
SASSARI

# ROBUST MULTI-MODAL AND MULTI-UNIT FEATURE LEVEL FUSION OF FACE AND IRIS BIOMETRICS

*Ajita Rattani and Massimo Tistarelli*


University of Sassari
Computer Vision Laboratory
Porto Conte Ricerche - Loc. Tramariglio
07041 Alghero (SS)
www.uniss.it

December 16, 2008

Technical Report No. CVL -2008-003

**Abstract**

Multi-biometrics has recently emerged as a mean of more robust and efficient personal verification and identification. Exploiting information from multiple sources at various levels i.e., feature, score, rank or decision, the false acceptance and rejection rates can be considerably reduced. Among all, feature level fusion is relatively an understudied problem. This paper addresses the feature level fusion for multi-modal and multi-unit sources of information. For multi-modal fusion the face and iris biometric traits are considered, while the multi-unit fusion is applied to merge the data from the left and right iris images. The proposed approach computes the SIFT features from both biometric sources, either multi-modal or multi-unit. For each source, the extracted SIFT features are selected via spatial sampling. Then these selected features are finally concatenated together into a single feature super-vector using serial fusion. This concatenated feature vector is used to perform classification.

Experimental results from face and iris standard biometric databases are presented. The reported results clearly show the performance improvements in classification obtained by applying feature level fusion for both multi-modal and multi-unit biometrics in comparison to uni-modal classification and score level fusion.

Keywords: Biometrics, Multi-modal data fusion, and Multi-unit data fusion, Face recognition, Iris recognition

# 1 Introduction

Biometrics refers to the use of physiological, biological or behavioural characteristics to establish the identity of an individual. These characteristics are unique to each individual and remain partially un-altered during the individual's lifetime [9].

In the recent years biometric authentication has gained a considerable improvement in both reliability and accuracy. Nonetheless, the best biometric systems to date present several drawbacks, such as limited applicability, vulnerability to spoofing attacks, less discriminant features, performance degradation due to noisy data, and others. Some of these limitations are inherent in the sensor technology or in the nature of the biometric trait itself. Consequently, a mono-modal biometric system rarely can cope with the variety of requirements in real applications. This is especially true in non-ideal scenarios, like outdoor environments, or highly demanding applications, such as in large-scale systems. By grouping evidence from multiple sources of information, multi-biometric systems [7] allow to overcome some of the drawbacks of the mono-modal biometric systems. Since the combination of multiple traits provides a better population coverage, multi-biometric systems represent the best solution to build a biometric platform with a wider usability. Multi-biometrics also offers an efficient counter-measure to spoofing, because it would be difficult for an impostor to simultaneously spoof multiple biometric traits of a genuine user [7].

Multi-biometric systems can rely on multi-modalities like face and fingerprint, multiple units like two or more fingerprints, multiple sensors like optical and capacitive

sensors, multiple representations like LDA and PCA features for face [7]. The resulting multiple sources of information can be integrated at various levels. Ross and Jain [13]presented a wide overview of multi-modal biometric solutions describing different levels at which fusion can be performed, i.e. sensor level, feature extraction level, matching score level and decision level. As reported in the literature [13] a biometric system that integrates information at an earlier stage of processing is expected to provide better performances than systems that integrate information at a later stage, because of the availability of more and richer information. Therefore, fusion at sensor and feature level is expected to be more efficient than fusion at matching score, rank and decision levels [4][3][6]. Despite of the abundance of research papers related to multi-biometrics, fusion at feature level is a relatively understudied problem [12][21][15]. One possible reason is that fusion at feature level is relatively difficult to achieve in practice because different modalities may require incompatible feature sets and the correspondence among different feature spaces may be unknown. Moreover, fusing feature vectors may lead to the problem of the curse of dimensionality: due to the large dimensions of the fused feature vector, a very complex matcher is required to correctly classify the data. Nevertheless, the fused feature vector may contain noisy or redundant data, thus leading to a degradation in the performance of the classifier [12].

Among many biometric traits, face and iris show a number of useful features for recognition. Face sensing is non-invasive and friendly. Moreover, it is relatively easy to acquire face images of good quality. On the other hand, iris-based identification systems proved to be among the best performing in many application domains. Nonetheless, both face and iris have their own limitations. The accuracy of face recognition is often affected by illumination, pose, shelter and facial expression, while current iris recognition systems require the active cooperation of the user. In a sense the two modalities are highly complementary. In fact, while face data is easy to acquire but delivers low to medium performances, iris data is more difficult to acquire but delivers very high performances. Moreover, as they both belong to the same biological trait (the head) they can be easily sampled at the same time and with the same sensing technology. Therefore, it is expected that fusion of face and iris may overcome the intrinsic limitations of each modality, while delivering the following advantages [19][20]:

1. The total error rate (the combination of the false accept rate and the false reject rate) is decreased with respect to the best of the two modalities.

2. The probability of successful spoofing attacks is reduced, because of the increased difficulty in manufacturing (or reproducing in a simulated video) an artificial head with both the face and irises faithfully reproducing the genuine user's appearance. On the contrary, it makes more difficult for an individual, in a surveillance scenario, to hide his/her identity.

3. The population coverage is larger than for a uni-modal biometric system. The adoption of the face modality can allow people with imperfect iris images to enroll, reducing the enrolment failure rate.

4. Conversely to other multi-biometric systems, which require multiple sensors, the combination of face and iris allows for simultaneous acquisition of the data with

a single device. Therefore, multi-modal and multi-unit fusion can be performed without introducing any additional cost for the hardware.

Few examples of face and iris fusion have been reported in the literature. The combination of iris and face at score level have been reported in [19][20]. References [16][5] are the only one in literature, in our knowledge, describing a methodology for feature level fusion. Both of them have applied projection to lower dimensional space for feature fusion and reduction using prior training and have used nearest neighbour distance for classification. This paper proposes a new approach for feature level fusion of face and both left and right iris from the same individual. In the proposed system, SIFT features are extracted from the face image and both irises. A spatial sampling technique is applied to select a subset of SIFT features individually from features extracted from each of the three biometric traits. The selected features are then combined into a single, higher dimensional fused feature set. Both the multi-modal (face and iris) and multi-unit (left and right irises) sources of information are combined to obtain a single super SIFT feature vector set which is used for matching and classification.

The main novelty of the work stems from the introduction of SIFT [10][11][17] features to process iris images and the combination of the features in a global SIFT feature vector set. The proposed approach has the advantage over [16][5] of adopting a feature extractor which delivers more stable features. The applied feature reduction technique is very simple and does not require a preliminary training. Also for classification, a direct matching technique for the fused feature vectors is adopted. Moreover, the work also presented the fusion of both multi-modal and multi-unit sources at feature level which further enhanced the results.

Experimental results on the database, composed of a subset of iris images from the CASIA version 3 database [1] and its chimerical combination with the Equinox database [2] are reported. The results obtained demonstrate the performance enhancement of feature level fusion in comparison to score level fusion and to the uni-modal systems. Section 2, describes the scale invariant features transform and its application for face and iris. Section 3, describes the feature level fusion of face with iris and also the left and right iris. Experimental results are presented in section 4. Conclusions are drawn in Section 5.

## 2 Scale Invariant Feature Transform *(SIFT)*

David Lowe proposed a method to extract distinctive and invariant features from images applied to general 3D object recognition [10]. This kind of features (also known as Scale Invariant Feature Transform or SIFT features) are invariant to image scale and rotation. They provide a basis for object representation to perform a robust matching across a substantial range of affine distortion, change in 3D viewpoint, addition of noise, and change in illumination. SIFT features are well localized in both the spatial and frequency domain, reducing the probability of disruption due to occlusion, clutter, or noise. In addition, SIFT features are highly distinctive. This property allows a single feature to be correctly matched with high probability against a large database of features, providing a basis for robust recognition. The cost of extracting these features

is minimized by adopting a cascade filtering approach (Fig. 1), in which the more time consuming operations are applied only at image locations which have been selected at the initial phase. The main computational steps applied to generate the set of image features are the following: David Lowe proposed a method to extract distinctive and invariant features from images applied to general 3D object recognition [10]. This kind of features (also known as Scale Invariant Feature Transform or SIFT features) are invariant to image scale and rotation. They provide a basis for object representation to perform a robust matching across a substantial range of affine distortion, change in 3D viewpoint, addition of noise, and change in illumination. SIFT features are well localized in both the spatial and frequency domain, reducing the probability of disruption due to occlusion, clutter, or noise. In addition, SIFT features are highly distinctive. This property allows a single feature to be correctly matched with high probability against a large database of features, providing a basis for robust recognition. The cost of extracting these features is minimized by adopting a cascade filtering approach (Fig. 1), in which the more time consuming operations are applied only at image locations which have been selected at the initial phase. The main computational steps applied to generate the set of image features are the following:

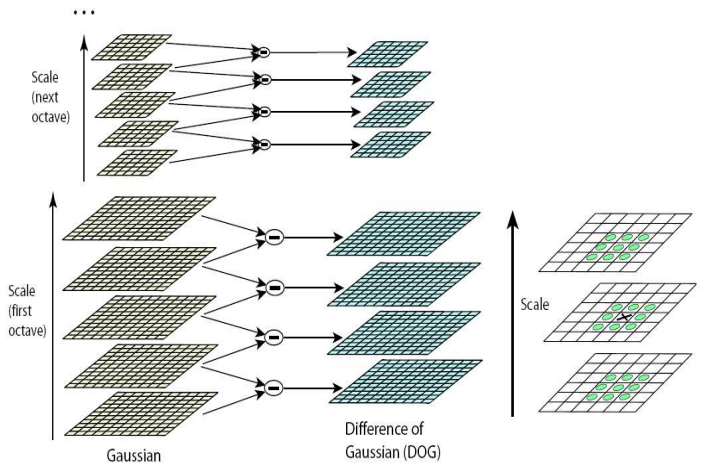


Figure 1: Computational schema for the extraction of SIFT features. Local extrema (marked with "X" on the right side of the figure) are extracted by comparing each pixel against its 26 adjacent pixels, within a $3 \times 3 \times 3$ neighbourhood spanning adjacent DoG images.

Scale-space extrema detection. The first processing stage searches over all scales and image locations. It is implemented efficiently by using a difference-of-Gaussian (DoG) function to identify potential interest points that are invariant to scale and orientation. Noting a Gaussian-blurred image as :

$$L(x, y, \sigma) = G(x, y, \sigma) \times I(x, y) \tag{1}$$

where $I(x, y)$ is the given image and $G(x, y, \sigma)$ is a variable scale Gaussian using 2.

The scale space of an image is then defined as a function of L(x,y,$\sigma$) using 1

$$G(x, y, \sigma) = \frac{\pi\sigma^2}{2}\ e^{-\frac{x^2+y^2}{\sigma^2}} \tag{2}$$

Stable keypoint locations in scale space can be detected by applying the method proposed in [10]. This method is based on the extraction of the scale-space extrema from the image filtered with a difference-of-Gaussian mask D(x,y,$\sigma$). The DoG filtered image can be computed from the difference of two Gaussians at nearby scales separated by a constant multiplicative factor $k$ using 3 and 4

$$D(x, y, \sigma, k) = (G(x, y, k\sigma) - G(x, y, \sigma)) \times I(x, y) \tag{3}$$

$$D(x, y, \sigma, k) = L(x, y, k\sigma) - L(x, y, \sigma) \tag{4}$$

Interest points (also called key points) are identified as local maxima or minima of the DoG filtered images across scales. The value of each pixel in the DoG filtered images is compared to its 8 neighbours at the same scale, plus the 9 corresponding neighbours at the lower and higher scales. If the pixel corresponds to a local maximum or minimum, it is selected as a candidate keypoint.

Accurate keypoint localization. At each candidate location, a detailed model fitting is applied to determine the correct keypoint location, scale, and the ratio of principal curvatures. This information allows rejecting points with local low contrast, which is sensitive to noise, or poorly localized along an edge. Orientation assignment. Based on the local image gradient directions, one or more orientations are assigned to each keypoint location. To determine the keypoint orientation, a gradient orientation histogram is computed in the neighbourhood of the keypoint, from the Gaussian filtered image at the closest scale to the keypoint's scale. The contribution of each neighbouring pixel is weighted by the gradient magnitude and a Gaussian window with $\sigma$ equal to 1.5 times the scale of the keypoint. Peaks in the histogram correspond to dominant orientations. A separate keypoint is created for the direction corresponding to the histogram maximum, and any other direction within 80% of the maximum value. All the properties of the keypoint are related to the keypoint orientation, to enforce invariance to rotation. For each image sample $L(x, y)$, at the chosen scale, the gradient magnitude $m(x, y)$ and the orientation $\theta(x, y)$ are computed using pixel differences using 5 and 6:

$$m(x, y) = \sqrt{L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1)^2} \tag{5}$$

$$\theta(x, y) = \tan^{-1}\left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}\right) \tag{6}$$

The orientation histogram is composed of 36 bins covering the 360 degree range of orientations with an approximation error equal to 10 degrees. Computation of the keypoint descriptors. The local image gradients are computed, at the selected scale, within a region around each keypoint. Once a keypoint orientation has been selected, the feature descriptor is computed as a set of orientation histograms, computed over a

$4\times4$ pixel neighbourhood. The orientation histograms are computed from the Gaussian filtered image which is closest in scale to the keypoint's scale. As for the previous stage, the contribution of each pixel is weighted by the gradient magnitude, and by a Gaussian with a $\sigma$ equal to 1.5 times the scale of the keypoint. Each descriptor contains an array of 4 histograms, each composed by 8 bins, computed around the keypoint. The final SIFT feature vector is composed of $4 \times 4 \times 8 = 128$ elements. In order to enforce the invariance to linear or affine changes in illumination, the feature vector is normalized. Due to the stability and robustness of these features, they have been recently applied to face and fingerprint [10][11][17]. Thus SIFT features are extracted from each face and iris image as explained above (Fig. 2). Each extracted SIFT features can be defined as $S = \{s_1, s_2, \ldots s_m\}$, where each feature $s_i = (x, y, \theta, \text{Keydesc})$ includes the spatial location (x, y), the local orientation $\theta$ and the key descriptor of size $1 \times 128$. Key descriptor i.e., Keydesc part of each SIFT feature is only invariant to affine transformations and contains large part of information about each SIFT point so it is only considered i.e, for each extracted SIFT point Keydesc is only used for experiments.

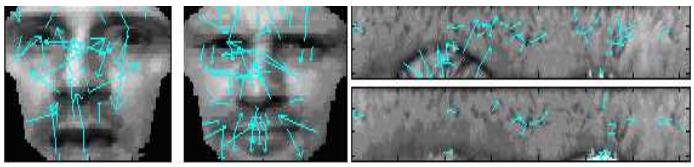

Figure 2: Extracted SIFT Features from the face and iris images. Even with the strong intra-class variations many common SIFT Features can be easily noticed.

# 3 Feature Level Fusion

Feature level fusion is performed by a simple concatenation of the feature sets obtained from different sources. Let $X = (x_1, x_2, \therefore x_m)$ and $Y = (y_1, y_2, \therefore y_m)$ denote the feature vectors extracted from two different sources. The vector $Z = (z_1, z_2, \therefore, z_m)$ is formed by the concatenation of the two feature sets $X$ and $Y$. Before performing the concatenation, the two vectors $X$ and $Y$ are normalized to ensure the same range and scales of values. In order to reduce the dimensionality of the feature space, a feature selection/reduction technique is also applied to the resultant feature vectors. The similarity between two individuals is determined by matching the instances of the vector $Z$ from the two individuals. The matcher simply computes the proximity between the two concatenated feature vectors $Z$ from the two subjects [12]. By applying SIFT features to represent both the face and iris traits, the entire process is greatly simplified without compromising recognition ability. In fact, the feature normalization process is not required, because the features from both sets are already commensurable, while the matching step is reduced to the selection of the nearest neighboring feature and the computation of the Euclidean distance metric. The entire matching process is described by the following steps:

1. Image Pre-processing and Feature Extraction: Both face and iris images are preprocessed to extract only the relevant part of the image containing useful information. A face template mask is applied to crop the central region of the face. The mask is registered and scaled to each face image on the basis of the position of the eyes and the face portion is cropped. In the reported experiments the eyes position are manually extracted but can be equally well extracted using a eye position extractor. It is worth noting that no further preprocessing is applied to the face images to compensate for illumination changes between samples. We deliberately choose to not perform an intensity normalization of the images to avoid a further processing step and also to fully validate the robustness of the SIFT-based representation. The iris images are segmented by using the technique proposed by Ross et al. [14] which is based on the development of geodesic curves. SIFT features are extracted from the segmented and normalized iris images. As a result of this step a set of SIFT features is obtained from the face image and the images containing the two irises from each subject.

2. The extracted SIFT features are then selected from both biometric traits by (Fig. 3)

   (a) dividing each image (face, left and right iris) into small windows of size $3 \times 2$ for the iris and $5 \times 5$ for the face. The total number of windows is equal to 256 for the face and 3780 for the each iris image. The window size has been determined after several trials and taking the size which best captures the information content in the data set. After performing several tests on different data, it was established experimentally that the optimal size only depends on the size of the input image.

   (b) For each window the SIFT descriptor with average minimum Euclidean distance from all other descriptors within the same window, is selected. In case there is only one descriptor in the window, it is selected for inclusion into the fused feature vector.

3. Feature Concatenation: The selected SIFT Features from the face and the two irises are combined into a single super feature vector set of dimension $N \times 128$, with $N$ equal to the total number of features selected from different trait.

$$\mathrm{F}_{(\mathrm{n+m+p})\times 128}$$

   where $F$ is the super fused feature vector which $n$ are the number of SIFT features belonging to face, $m$ belongs to left Iris and $p$ are the features belonging to right iris image.

4. Feature Matching: The fused feature sets from two individuals are matched by comparing each feature element based on the associated descriptors. In order to match two fused feature sets $f_1$ and $f_2$ . Given a feature element $p_{11}$ in $f_1$ , the distances $d_1$ and $d_2$ between the first closest element $p_{11}$ and the second closest element $p_{21}$ in $f_2$ are computed. If the ratio is below 60%, then the vector

element $p_{11}$ is tagged to match with element $p_{21}$. The matching score between the two fused feature vectors is equal to the number of matching elements [17]. It is worth noting that this procedure maximizes the separation between matching and non-matching vector elements, thus reducing the probability of false matches between vector pairs.

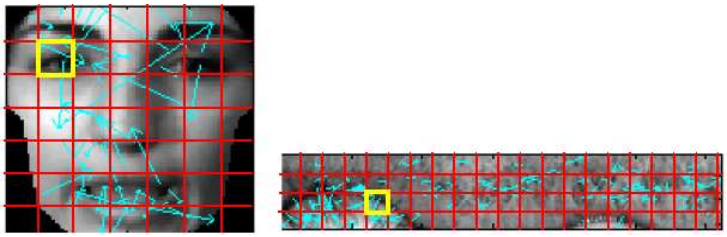

Figure 3: Example of feature selection by partitioning the face and iris images into a set of small windows and selecting one SIFT feature per window. The yellow boxes show an example of more than one SIFT feature in a single partition. Windows with no features are ignored.

# 4 Experimental Evaluation

The CASIA iris database version 3 [1] and the Equinox face database [2] have been used to evaluate the performance of the proposed system. A total of 57 subjects with ten instances per each client have been taken from the Equinox database. The left and right normalized iris images of 125 clients from the CASIA ver.3 database have been used for experiments.

## 4.1 Protocol for Performance Evaluation

The following evaluation procedure has been applied for mono-modal and multi-biometric matching, with $N$ subjects and processing 10 samples for each biometric trait per subject:

Training: one template is built for each subject using one image for each modality, i.e. one face image and one image for each of the two irises. The SIFT features are extracted from each image and both uni-modal and multi-modal feature vectors are built for the face, irises and the combination of the two irises and the face with one or two irises. In the performed experiments, the matching scores were computed to determine the discrimination capability of the single and fused representation. For this reason it was not necessary to train a multi-modal classifier to perform the final classification among the subjects in the database.

Testing: Nine samples per person are used for testing and to generate the client matching scores. For the mono-modal systems, the impostor scores are generated by matching the representation of each client against the first sample of all the other individuals in the dataset. In case of multimodal testing the client is tested against the

first face and iris samples of the remaining of the chimerical users. Therefore, in total $9 \times N$ client scores and $9 \times N \times (N-1)$ impostor scores for each of the uni-modal and multimodal representations are generated.

## 4.2 Multi-unit Fusion: Left and Right Irises

For testing the application of SIFT representation to the iris biometric, 125 clients from the CASIA version 3 database [1] with ten left and ten right iris images were used as mentioned in 4.1. The iris images are normalized and segmented as discussed in [14]. Left and Right iris are confirmed to contain discriminatory information for each individual so their combination is a source of complementary information [18]. Thus feature level fusion of left and right iris using SIFT is performed as described in section 3 and it was evaluated against left and right uni-modal traits. The performance is evaluated as proposed in section 4.1. In total, $125 \times 9 = 1125$ client scores and $125 \times 124 \times 9 = 139500$ impostor scores are generated for each iris. The ROC curves obtained from the matching scores of the left and right iris images using SIFT Features and feature level fusion are shown in Fig. 4. As it can be noted from the ROC curves, the feature level fusion of these complementary information sources considerably enhances the identification capabilities.

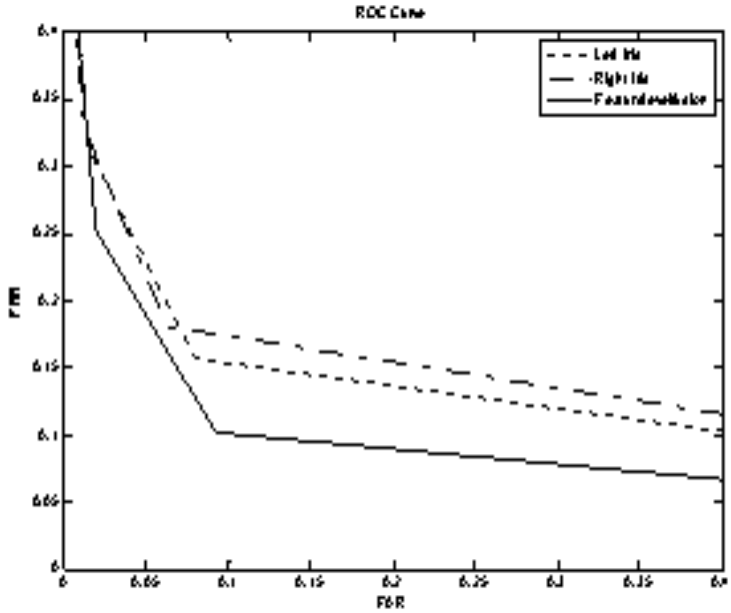


Figure 4: ROC curves representing the matching performance of individuals based on the SIFT representation extracted from the left and right iris and the feature level fusion of the two irises.

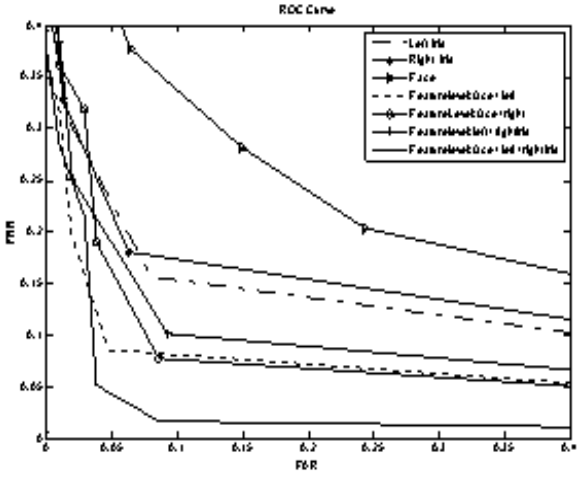


Figure 5: ROC curves representing the matching performance of individuals based on the SIFT representation extracted from the face, the left and the right iris and the feature level fusion of the three traits.

### 4.3 Multi-modal with Multi-unit Fusion: Face, Left and Right Irises

To test the performance of matching the fused SIFT representation based on face and iris, a chimerical database is obtained by combining the iris samples with 57 clients from the Equinox face database [2]. The SIFT features are extracted from the normalized face and iris images. The SIFT representation of the face and iris biometric traits are fused at feature level as detailed in section 3. The performance is evaluated as reported in section 4.1. In total $57 \times 9 = 513$ client scores and $57 \times 56 \times 9 = 28728$ impostor scores are generated for the uni-modal face matching and the face and iris fused vector matching. The ROC curves representing the error distributions of the feature level fused matching are shown in Fig 5. The curves were generated from the scores obtained by matching the mono-modal face representations and the multi-modal feature level fusion with left and right iris images. It is worth noting that the error distributions are lowered by combining more information at the feature level. The performance of the feature level fusion of face and the two irises is compared with the score level fusion in Fig 6. As it can be noted, the data fusion performed at the earlier stage, i.e. feature level, produces better performances in terms of error distributions than the fusion at the score level.

## 5 Conclusion

The fusion of face and iris biometrics is a natural process in a multi-biometric system which, in principle, can be implemented far more easily than other biometric fusion systems. At the same time, the two biometrics are strongly complementary. In fact, while face biometric samples are easier to acquire with a conventional camera, iris demonstrated superior performances with very low false acceptance rates. This suggests

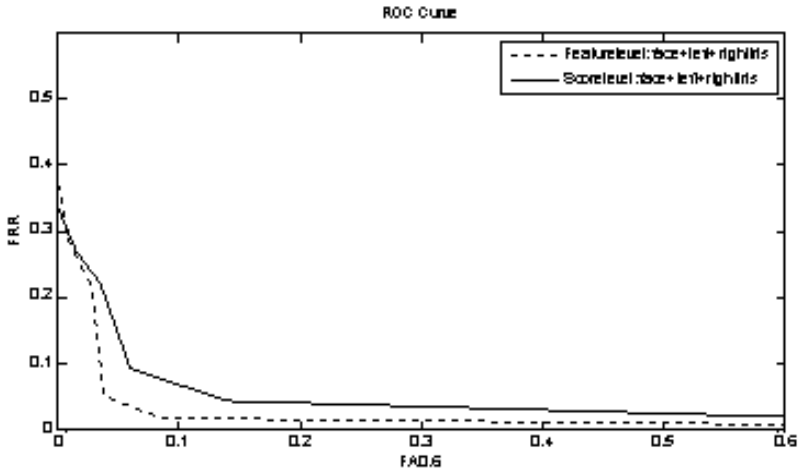


Figure 6: Comparison of the feature level fusion, of the SIFT representations of the face, the left and the right irises, with the score level fusion of the matching scores of the individual modalities.

an image acquisition system with a camera acquiring face images at a high resolution to provide shape and textural information on both the two irises and the face itself. Current camera devices allow to sample up to $15M$ pixels images at a rate of 30 frames per second and deliver the data stream over a fast Ethernet channel. By properly controlling and triggering the image acquisition with a fast face detection algorithm to keep the head within one third of the total image area, the resulting images can provide iris images up to $64K$ pixels. Several systems to acquire iris from a distance have been proposed and remarkably demonstrated the possibility to acquire both the face and the iris at the same time [8].

In this paper a novel approach to feature level fusion of face and iris has been proposed. The main advantages of the proposed fusion method are the ease of implementation and the robustness of the resulting representation. Due to the scale-space analysis, SIFT features proved to be very stable and almost insensitive to illumination variations while providing a scale and translation invariant representations. At the same time, the adoption of a common feature representation greatly simplifies the normalization and concatenation processes in the feature fusion process. From the experiments performed on a mixed database obtained combining face images from the Equinox database [2] and iris images from the CASIA v. 3 database [1] the representation based on feature level fusion demonstrate superior matching performance with respect to all the uni-modal and the single iris and face based representations. Several issues are still under investigation, such as the optimal choice of the face and iris sub-sampling for feature selection. The adoption of clustering techniques can be also investigated.

# References

[1] In *http://www.cbsr.ia.ac.cn/english/IrisDatabase.asp*.

[2] In *http://www.equinoxsensors.com/products/HID.html*.

[3] S. Fischer B. Duc, G. Maitre and J. Bigun. Person authentication by fusing face and speech information. In *In Proc. of the AVBPA*, volume LNCS. Springer Verlag, 1997.

[4] J.S. Mason C.C. Chibelushi and F. Deravi. Integration of acoustic and visual speech for speaker recognition. In *In EUROSPEECH'93s*, pages 157–160, 1993.

[5] J. Gan and Y. Liang. A method for face and iris feature fusion in identity authentication. In *IJCSNS*, volume 6(2B), 2006.

[6] L. Hong and A. Jain. Integrating faces and fingerprints for personal identification. In *IEEE Transaction on PAMI*, volume 20(12), pages 1295–1307, December 1998.

[7] A. K. Jain and A. Ross. Multi-biometric systems. In *Communications of the ACM*, volume 47(1), pages 34–40, 2004.

[8] K. Hanna R. Kolczynski D.J. LoIacono S. Mangru M. Tinker T.M. Zappia W.Y. Zhao J.R. Matey, O. Naroditsky. Iris on the move: Acquisition of images for iris recognition in less constrained environments. In *Proc. of the IEEE*, volume 94 (11), pages 1936–1947, 2006.

[9] A. Jain L. Hong and S. Pankanti. Can multi-biometrics improve performance. In *Proc. of AutoID*, pages 59–64, 1999.

[10] Lowe and G. David. Object recognition from local scale invariant features. In *International Conference on Computer Vision*, pages 1150–1157, 1999.

[11] E. Grosso M. Bicego, A. Lagorio and M. Tistarelli. On the use of sift features for face authentication. In *Proc. of Int Workshop on Biometrics in association with CVPR*, 2006.

[12] A. Ross and R. Govindarajan. Feature level fusion using hand and face biometrics. In *Proc. of SPIE Conference on Biometric Technology for Human Identification II*, pages 196–204, 2005.

[13] A. Ross and A.K. Jain. Information fusion in biometrics. In *Pattern Recognition Letters*, volume 24, pages 2115–2125, 2003.

[14] A. Ross and S. Shah. Segmenting non-ideal irises using geodesic active contours. In *Proc. of Biometrics Symposium (BSYM), (Baltimore, USA)*, 2006.

[15] G. Gyaourova S. Singh and I. Pavlidisu. Infrared and visible image fusion for face recognition. In *SPIE Defense and Security Symposium*, pages 585–596, 2004.

[16] B. Son and Y. Lee. Biometric authentication system using reduced joint feature vector of iris and face. In *AVBPA*, pages 513–522, 2005.

[17] S. Pankanti U. Park and A. K. Jain. Fingerprint verification using sift features. In *Proc. of SPIE Defense and Security Symposium*, 2008.

[18] D. Zhang X. Wu, K. Wang and N. Qi. Combining left and right irises for personal authentication. In *Energy Minimization Methods in Computer Vision and Pattern Recognition*, volume LNCS Springer 4679, pages 145–152, 2007.

[19] T. Tan Y. Wang and A. K. Jain. Combining face and iris biometrics for identity verification. In *Proc. of 4th Int' Conf. on Audio- and Video-Based Biometric Person Authentication (AVBPA)*, pages 805–813, 2003.

[20] K. Pan Stan Z. Li Z. Zhang, R. Wang and P. Zhang. Fusion of near infrared face and iris biometrics. In *ICB*, 2007.

[21] X. Zhou and B. Bhanu. fusion of face and gait for human recognition at a distance in video. In *International Conference on Pattern Recognition*, 2006.