Detection Optimization for Biometric Applications with Non-Linear Classification Models

Sorin Soviany[1], Cristina Soviany[2], Sorin Puşcoci[3], Mariana Jurian[4]

[1]National Communication Research Institute (I.N.S.C.C.), Romania

[2]DES Technologies, Belgium

[3]National Communication Research Institute (I.N.S.C.C.), Romania

[4]University of Pitesti, Romania

[1]sorin.soviany@inscc.ro; [2]cristina.soviany@ides-technologies.com; [3]sorp2006@gmail.com; [4]mariana.jurian@yahoo.com

***Abstract*-The paper proposes a classification method for people identification accuracy improvement in which the biometric system is trained not for all enrolled individuals but only for a few target identities to be recognized, therefore reducing the computational complexity for the large-scale biometric identification. The biometric detectors are relying on non-linear models which are more suitable for the real biometric data with high degree of intra-class variance; therefore they improve the people recognition accuracy even for the most difficult cases.**

Keywords- Detector; Identification; Non-Linear

I. INTRODUCTION

The actual biometric systems still exhibit performance issues if the people recognition uses only their biometric patterns, without any additional identificators. Most of the actual accuracy improvements were achieved for verification applications (in which the decision is to validate or invalidate a claimed identity). For instance, Jain and Ross evaluated a multimodal biometric system with fingerprint, face and hand geometry using the weighted sum fusion rule; they achieved a Genuine Acceptance Rate of more than 98.5%, for the identity verification only [8]. The actual developments are searching for innovative security biometric solutions with a suitable trade-off cost-performance in identification. This trade-off is achieved by feature space dimensionality reduction or by choosing the optimal classifiers according to the desired performance level [1]. The multi-class problem of the biometric identification is still a big challenge for biometric security systems design; models like Support Vector Machines provided significant improvements especially for 2-class problems [10] (i.e. biometric verification systems with genuine/impostor decisions but without precisely identification of the subjects).

In this paper we proposed a method for data classification focused on identification biometric applications. The system is trained not for all enrolled persons but only for a few target identities. This approach enables a significant computational complexity reduction which is useful for large-scale identification biometric applications. The main application is the remote access control to medical databases; this application requires an efficient authentication solution, with an optimal adjustment of the provided security level, according to the user's authorization degree within the telemedicine system. The remainder of this paper is structured as follows. In section 2 we present the system functional architecture. Section 3 presents the applied techniques for feature generation. In section 4 the proposed classification method is detailed. The experimental results are presented and discussed in section 5. Finally section 6 concludes our research.

II. SYSTEM ARCHITECTURE AND AVAILABLE DATASETS

We applied the proposed identities detection method on a multimodal biometric security system with 2 biometric: fingerprint and palmprint. The security system functional architecture is depicted in Fig. 1. This solution is designed for a medical database access security system providing a reliable remote access control for the authorized users. Also the designed system performs on a reduced-sized feature space (less than 10 features for each biometric).

The system functions are performed in 2 stages: *pre-processing* for templates generation and *processing* for identification. In the 1st stage the biometric templates are generated for enrollment and for further recognition. In the 2nd stage the individual's recognition is performed based on an innovative classification model in which the biometric data matching is relying on a special kind of classifiers called detectors. Finally the *application-level decision* has to allow or deny the access to the protected resource (the medical database) according to the authorization degree.

The available biometric data are provided from 20 users of the medical database. The fingerprint and palmprint biometric templates are generated from 5 images per person. The biometric datasets are randomly divided in 2 subsets: a training subset for the classifiers design and a validation subset which for the identification performance assessment. These 2 data subsets should be independent. We did not apply any cross-validation technique like K-fold or leave-one-out.

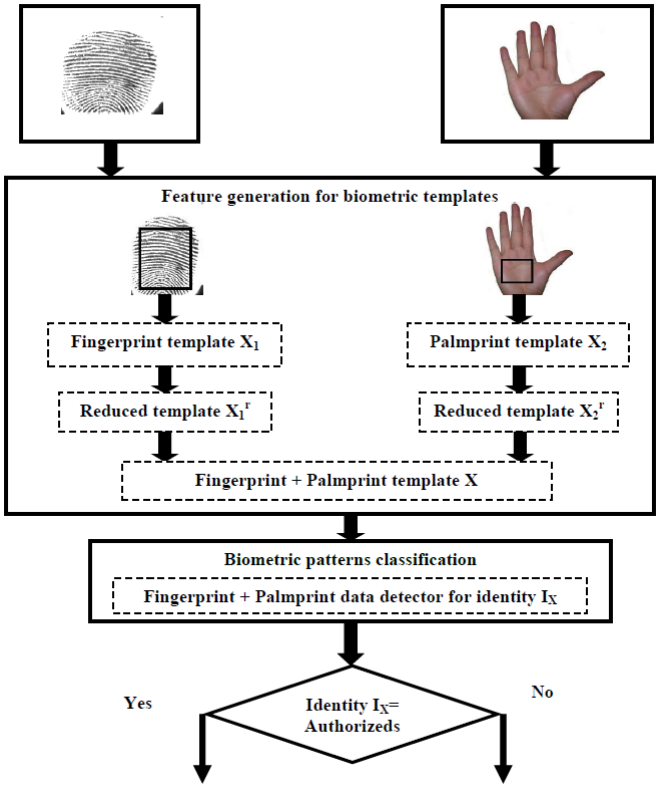

Fig. 1 The basic functional architecture for the security system

III. FEATURE GENERATION

In this stage the designed system performs the following operations: selecting ROIs (region of interest) in fingerprint and palmprint images, feature extraction from the selected ROI, feature space dimensionality reduction for each biometric and finally feature-level fusion. The final result is a single biometric template for each person, putting together fingerprint and palmprint features.

A. Regions of interest (ROI)

In biometric applications data (fingerprint and palmprint images in our application), the regions of interest are selected to provide only the relevant part of the images for feature extraction, without any background not useful for the biometric recognition. For both biometrics we used an automatic ROI selector in which the biometric ROIs are resulting from the classifier decisions. First we manually selected a rectangular region in a master (high-quality) image for each biometric (Fig. 2). Then we trained 2 detectors for the selected fingerprint and palmprint ROI. We applied the Fisher classifier for fingerprint and palmprint ROI detection because it had the lowest classification error rates even while using low-sized training image datasets.

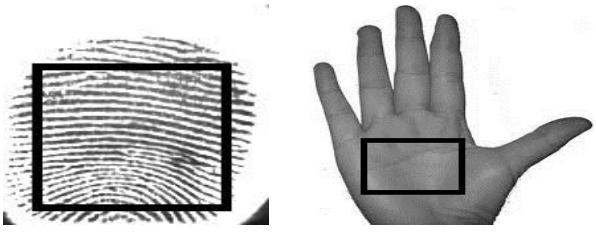

Fig. 2 ROIs in fingerprint and palmprint images for feature extraction

Therefore we did not need too many images to design efficient detectors for biometric ROIs further providing optimal features to perform the individual recognition. For 5 selected images per biometric per person, the fingerprint ROI detector provided an average classification error rate of 0.065 (Fig. 3a); for the palmprint ROI detector the average error rate was 0.11 (Fig. 3b). We achieved these values by averaging on 10 experiments and also by averaging on the 2 classes (ROI and non-ROI).

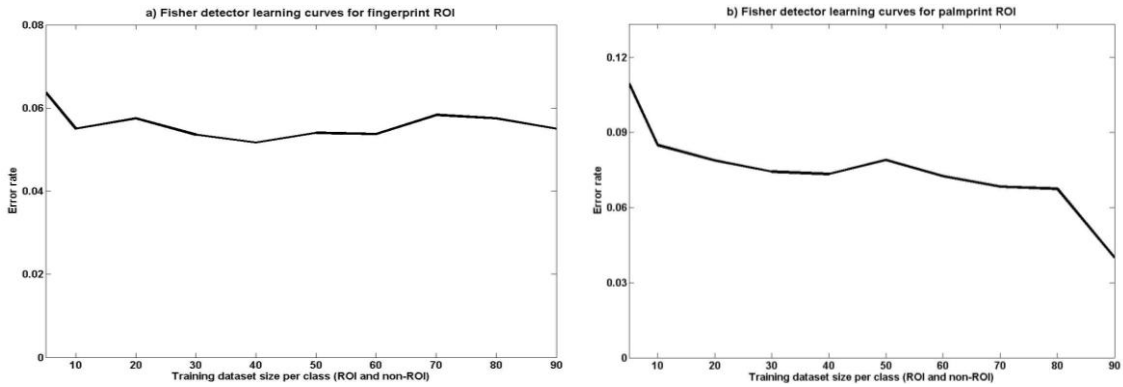

Fig. 3 ROIs Fisher detectors learning curves for: a) fingerprint; b) palmprint

However, as much as we are only interested in per ROI-class error minimization, finally we optimized the 2 detectors for fingerprint and palmprint ROI automatic selection by fixing their operating points in order to minimize the classification error rate on ROI class, and not for the per-class averaged error rate. This performance trade-off helps us to provide an optimal region for the further biometric feature extraction. We found out the following optimal operating points for the 2 ROI detectors: for fingerprint ROI detector an optimal op.point with an error rate on ROI class of 0.05 and for palmprint ROI detector an optimal op.point with an error rate on ROI class of 0.10.

B. Feature extraction

For fingerprint and palmprint feature extraction we applied the co-occurrence matrices in order to better exploit the textural features from the selected images. Each element of the co-occurrence matrix estimates the probability of a certain gray-level for a given pixel while a displaced pixel has another gray-level, according to eq. (1) [2, 3, and 4].

$$C_{\Delta_x,\Delta_y}(i,j) = P\left\{I(x,y) = i, I(x+\Delta_x, y+\Delta_y) = j\right\} \tag{1}$$

in which

C is the co-occurrence matrix computed for one m x n image *I*;

Δ_x, Δ_y are the displacement values.

The resulting feature space size is 16 for fingerprint data and 25 for palmprint data.

C. Dimensionality reduction

Another requirement for the system design is the feature space size reducing to less than 10 for each biometric. For the further classification stage more features need more training biometric templates to provide the desired performance and this is not very convenient for most of the biometric application users. On the other hand a high-sized feature space does not mandatory involve more accuracy in people recognition, as much as not all the extracted features have the same discriminant value. Finally, as much as the last operation of the pre-processing stage is feature-level fusion by fingerprint and palmprint templates concatenation, the feature space reduction for each biometric is required to ensure a low computational complexity of the classification stage.

We performed this operation by transforming the input templates with some projections (PCA-Principal Component Analysis) and LDA (Linear Discriminant Analysis); then we applied a feature selection procedure for each biometric in order to maintain only the best features. PCA is an unsupervised projection transform which provides the most variant features from the initial set, but without considering their class membership [5]. LDA projects the input data preserving their class membership [6]. For the selection step we applied the forward-search procedure and the criterion was 1-NN (nearest-neighbor rule) classification error rate because of its main property to limit the classification error rate [7]:

$$\varepsilon^* \le \varepsilon_{1-NN} \le 2\varepsilon^* \cdot (1-\varepsilon^*) \le 2\varepsilon^* \tag{2}$$

where ε_{1-NN} is the error rate for the 1-NN classifier and ε^* is the optimal Bayesian classifier error rate. Finally the resulting feature spaces sizes are the following: 7 features for fingerprint data and 9 features for palmprint data.

D. Feature-level fusion

In the feature-level fusion step we applied the simplest scheme which is feature vectors concatenation, according to Jain

and Ross [8]. This feature biometric fusion is convenient because it does not require many constraints with respect to the compatibility among the fused features. The main drawback is that the increasing dimensionality usually leads to the well-known problem called coarse of dimensionality [6]. In order to prevent or to limit this effect, we previously reduced the features number for each of the 2 biometrics (fingerprint and palmprint). Therefore we could apply this simple concatenation-based feature-level biometric data fusion (or pre-classification biometric fusion).

IV. BIOMETRIC DETECTION

A. The detection principle. Significance for the identification biometric application

A detector is a special kind of classifier which is trained for only one target class, thereby neglecting all other classes of the problem [9]. The detection is achieved by the model output thresholding on the target class. As much as our biometric application is an identification one and also not all the medical database users have the same authorization degree, we could apply detectors for finding the target identities. We are handling the biometric identification as a multi-class problem in which all the biometric templates for one person are representing a class, therefore each class corresponds to an identity. Particularly, a biometric detector is trained for a certain person to be recognized. The detection principle for the biometric identification is the following: if I is the target person identity, then the classification (identification) rule is

$$y(x) \geq \theta_I \Rightarrow Identity(x) = I \tag{3}$$

where: x is the biometric template which the unknown person applies to the system input in order to perform his/her identification; θ_I is the threshold for the identity detection and $y(x)$ is the underlying model output of the detector.

A main advantage of the detection approach for biometric recognition in people identification is the computational complexity reduction by handling a multi-class problem in a 2-class like problem or almost one-vs.-all approach; in this way, the underlying model has to compute the main parameters only for the target identity (class), therefore reducing the overall time for identification. The biometric detector is focused only for one target person and not for all the enrolled users of the protected resource.

B. Non-linear detectors models for biometric data

In many biometric applications there are challenges derived from the various sources of noise which are decreasing the biometric templates quality and finally the identification accuracy. The biometric templates often exhibit a certain degree of intra-class variance either due to lack of technical skills of persons in using the biometric devices or to some unsuitable environment conditions for data acquisition.

These challenges in accurate biometric systems design could be approached by using detectors in which the underlying models are optimized non-linear classifiers. Actually we applied a kernel SVM (Support Vector Machine) model for our detectors. The non-linear (kernel) SVM model is most suitable for this identification task because the available biometric data exhibit a high degree of non-linear separability even for the reduced feature space. In our application there are 3 persons with the highest authorization degree for the protected resource (medical database). Therefore we will design 3 detectors, each of them being trained for one identity recognition.

In this modeling I_1, I_2 and I_3 are representing the classes containing the biometric templates for the 3 selected identities. Each of the 3 detectors is designed starting to the following underlying model [6]:

$$g_k(x_k) = \text{sgn}\left(\sum_{i=1}^{N_{sk}} \lambda_{i_k} y_{i_k} \cdot K(x_{i_k}, x_k) + w_{0_k}\right), k = \overline{1,3} \tag{4}$$

in which:

x_{i_k} is a training biometric sample represented in the fused reduced feature space (fingerprint+palmprint) and applied as input for the designed identification system;

N_{s_k} is the support vectors number for the kernel SVM-based identity I_k detector;

λ_{i_k} are the Lagrange multipliers which are achieved for the optimization problem of finding the maxim boundary hyperplane;

y_{i_k} are the class labels (+1 for the target identity and -1 for all the other identities);

$K(x_{i_k}, x_k)$ is the kernel which we applied for our biometric data. This provides the non-linearity approach in handling the

biometric data detection. We selected a hyperbolic tangent function given by [6]

$$K(x_k, z_k) = \tanh(\alpha_k x_k^T \cdot z_k + \beta_k), k = \overline{1,3} \tag{5}$$

with the following parameters values that fit on our experimental data:

- for the 1st detector (identity I_1) $\alpha_1 = 1.5, \beta_1 = 0.8$;
- for the 2nd detector (identity I_2) $\alpha_2 = 1.8, \beta_2 = 1.2$;
- for the 3rd detector (identity I_3) $\alpha_3 = 2.0, \beta_3 = 1.0$

w_{0_k} is the offset parameters of the maxim boundary hyperplane and given by [6]

$$w_{0_k} = \frac{1}{N_{s_k}} \sum_{s_k \in V_{s_k}} \left[y_{s_k} - \sum_{j_k \in V_{s_k}} \lambda_{i_k} y_{j_k} \cdot K(x_{j_k}, x_{s_k}) \right] \tag{6}$$

where:

V_{s_k} is the support vectors set which are found for our biometric data detection with non-linear (kernel) SVM classifiers.

A SVM classification model (either linear or non-linear) output is not a statistical similarity (or matching) indicator neither a confidence level for the class membership (in our application the person identity). It is only a distance measure with respect to the separating hyperplane in the feature space [10]. In order to transform the SVM detector output into a class-posterior probability $P(I_k \mid x)$ (where I_k is the identity for which we are training the detection-based biometric system), we should apply a normalization technique based on the sigmoid function which provides a common numerical range between 0 and 1 for the classifier outputs. It provides a matching score according to

$$S_k = P(I_k \mid x) = \frac{1}{1 + \exp(-A_k \cdot g_k(x) - B_k)}, k = \overline{1,3} \tag{7}$$

where $g_k(x)$ is the SVM output for a test biometric sample *x* (i.e. the biometric pattern of the person to be recognized). Also the coefficients have the following best values for each of the 3 trained detectors:

- for I_1 detection: $A_1 = 2, B_1 = 1.5$;
- for I_2 detection: $A_2 = 1.5, B_2 = 1.5$;
- for I_3 detection: $A_3 = 2.5, B_3 = 2.0$

This normalization allows to threshold the SVM output classifiers to achieve the biometric detectors. Finally the identification decision rule becomes:

$$S_k \geq \theta_k \Rightarrow Identity(x) = I_k \tag{8}$$

where the threshold θ_k is fixed for the target identity based on the experimental data (the available training biometric templates).

C. Learning curves for the fingerprint and palmprint detectors

We performed the training and testing process in 10 experiments. Given the feature vectors (fingerprint+palmprint) with 16 components, we trained the biometric detectors by *grouping the available biometric data in 2 classes*: the 1st class I_k (k=1, 2 and 3 respectively) contains the focused (target) identity data and the 2nd class contains the biometric data for all the other (19) identities. Therefore in each experiment we performed 3 *training processes*, one for each of the 3 target identities. Then we evaluated the detectors performance on the independent *validation (testing) data subsets* which we randomly generated by splitting the original data sets. We repeated for 10 times this procedure (training for each identity and then testing for validation).

The optimal training biometric data set sizes are resulting from the non-linear (kernel) SVM-based detectors learning curves. We plot these curves (Fig. 4) for the 3 detectors and for 10 samples sized-steps of the training biometric

(fingerprint+palmprint) data sets. The error rates are averaged on 10 experiments and also they are computed by averaging over the target identity and all other non-target identities, according to:

$$IER(I_k) = \frac{1}{10} \cdot \frac{FN(I_k)}{TP(I_k) + FP(I_k)}, k = \overline{1,3} \quad (9)$$

where:

$IER(I_k)$ is the identification error rate for the identity I_k detector;

$FN(I_k)$ (False Negative on identity I_k) is the number of wrong decisions for a biometric pattern belonging to the true identity I_k (how many times a biometric pattern belonging to the person I_k is misclassified as belonging to any of the other enrolled persons);

$TP(I_k)$ (True Positive on identity I_k) is the number of correct decisions for the identity I_k (showing how many times, in the performed experiment, the biometric pattern belonging to the true identity I_k is assigned to this real identity);

$FP(I_k)$ (False Positive on identity I_k) is the number of wrong decisions on a biometric pattern belonging to the true identity non-I_k (i.e. how many times, in this experiment, the biometric data belonging to any of the other enrolled identities is misclassified as belonging to the target identity I_k);

In (9) the denominator evaluates the overall decisions on identity I_k.

Also these errors are evaluated on the independent validation datasets randomly extracted from the initial design data.

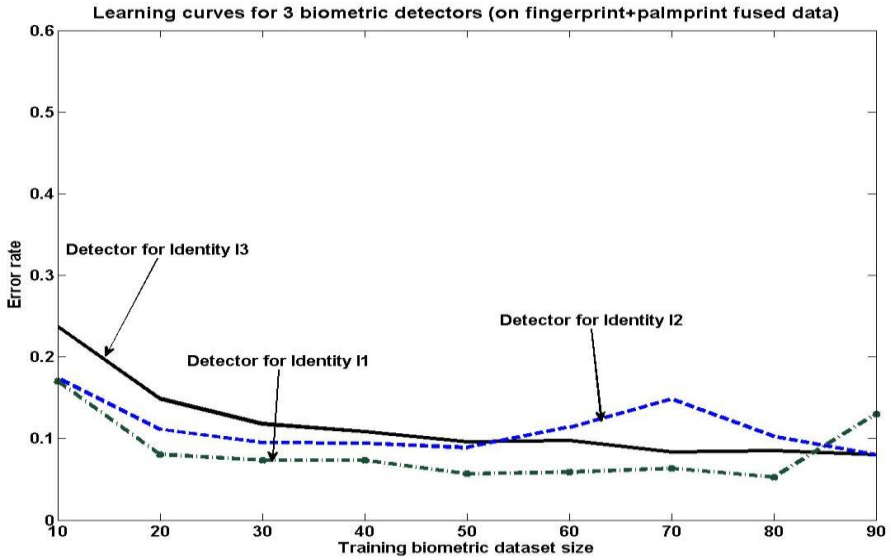

Fig. 4 Learning curves for the 3 identities trained detectors (IER)

We separately trained the detectors for fingerprint+palmprint biometric data, also adjusting the parameters in order to improve their behavior on these available data. The overall feature space size is 16 and it is resulted after the feature-level (or pre-classification) fusion stage. We depicted the curves which are corresponding to the best detectors behavior on the experimental data.

V. EXPERIMENTAL RESULTS. OPTIMIZATIONS

We evaluated the proposed method using MATLAB-based environments, actually PRT (Pattern Recognition Toolbox) and PerClass, respectively. A piece of PerClass source code which we used to trace the learning curves for the SVM-based trained detectors is the following:

```
[A1_tr, A1_ts]=randsubset (A1, 0.5)
step= [10:10:90];
err_SVM= [];
for i=1:length(step)
  t=randsubset (A1_tr, step (i))
```

```
p=sdsvc (t)
Pd=sddecide (p)
err_SVM (i) =sdtest (A1_ts, pd)
  end
```

PerClass is an environment for Pattern Recognition systems design which also is very suitable for optimization tasks in PR applications like biometric recognition [9]. One of its main advantages is that it allows calling the functions into run-time applications outside MATLAB.

A piece of the file containing the experimental data used for our evaluation is given here (actually these data are derived from one person fingerprint image using the textural approach for the feature extraction). The input data for the PerClass functions are actually matrices-based representation (arrays).

```
% data (16 features), lab

% import using: a=sdimport ('P1_A1_100.txt','skip', 2,'data', 1:16,'lab', 17)

0.00000000000000,0.08333333333333,0.00000000000000,0.08333333333333,0.16666666666667,0.33333333333333,0.00000000000000,0.33333333333333,0.00000000000000,  0.83333333333333,0.08333333333333,0.00000000000000,0.08333333333333,0.00000000000000,0.00000000000000,0.00000000000000, P1

0.16666666666667,0.16666666666667,0.00000000000000,0.16666666666667,0.50000000000000,0.00000000000000,0.00000000000000,0.00000000000000,  0.33333333333333,0.08333333333333,0.00000000000000,0.08333333333333,0.50000000000000,0.00000000000000,0.00000000000000,P1

0.50000000000000,0.16666666666667,0.00000000000000,0.16666666666667,0.16666666666667,0.00000000000000,0.00000000000000,0.00000000000000,0.25416666666667,0.01875000000000,0.00000000000000,0.25416666666667,0.01875000000000,0.00000000000000,0.01875000000000,0.02083333333333,0.02291666666667,0.00000000000000, P1
```

Actually the applied co-occurrence matrix method provides more than 1.000 feature vectors from only one image per person; finally we reduced this number by applying the ROI-based approach.

The first view on the biometric detectors performance for 3 target identities is resulting from their learning curves (Fig. 4) which are allowing to provide a better design, for instance by a suitable sizing of the required training biometric samples in order to enhance the focused persons identification accuracy. We could see that the optimal biometric templates needed for the 3 detectors training should be around 40 to 50 samples for the focused class (i.e. the target identity to be recognized). For the other identities (grouped in the detector non-target class) we will use less training sample, as much as the detector design is most focused on the target identity. Therefore in this biometric data classification system we provided a dataset (fused fingerprint+palmprint templates) containing 50 biometric samples for the focused identity. This is also suitable for a relatively small feature number (which is 16 in our experiment), knowing that more feature means more training samples to ensure a better coverage of the whole feature space.

Also from these learning curves we could notice that the 1st detector (which is designed for recognition of the most authorized user of the medical database) shows a better behavior on the available biometric data for most of the training set sizes.

So far the expected performance for the 3 people's identity detection, as resulting from the learning curves depicted in Fig. 4, should be described by the following values for the averaged identification error rates if we train the identities detectors with 50 biometric samples per target identity:

- for the 1st detector (identity I_1): $IER(I_1) = 0.06$;
- for the 2nd detector (identity I_2): $IER(I_2) = 0.09$;
- for the 3rd detector (identity I_3): $IER(I_3) = 0.1$

We achieved these values on the default thresholding of the normalized classifiers output and without any further optimization depending on the particular application requirements.

The additional optimization could be done on the detectors ROC curves by fixing the optimal thresholds in order to get the best operating points according to the security applications requirements (i.e. how precisely such be the person identification depending on the potential consequences of an unauthorized access to the protected database). Actually this optimization basically means a trade-off between the detectors performance on target and non-target identities. Also we can specify a rejection rate for the testing biometric data, as much as many identification errors are resulting from low-quality enrolling or

testing biometric templates. For instance, almost 5% from the overall fingerprints could not generate suitable biometric templates to be enrolled and/or used for authentication. This is why we applied a classification rejection rate of 5% from all biometric feature vectors which are used for system design.

The operating points for the 1st detector are represented on the ROC curve depicted in Fig. 5.

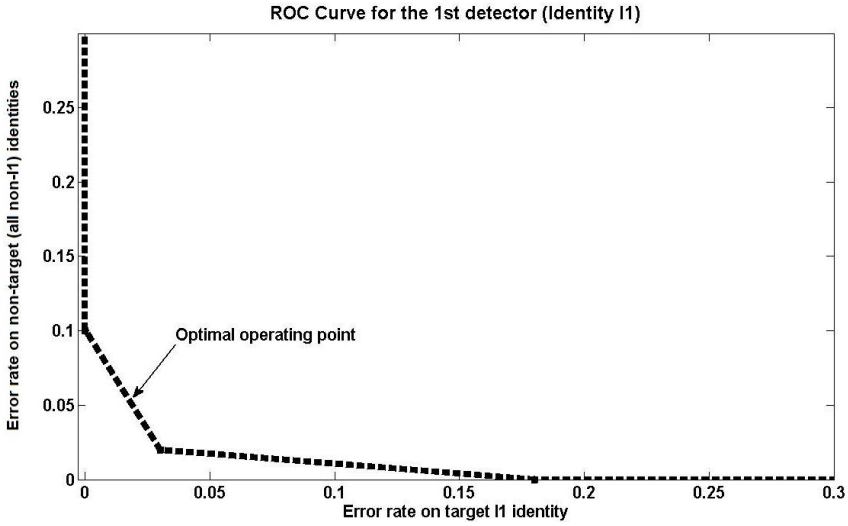

Fig. 5 Performance optimization for the 1st detector (Identity I_1)

For the 1st detector we fixed the optimal operating point by reducing the detection threshold to 0.50, instead of the default value of 0.70. This allows us to reduce the identification error rates on the target identity (I_1) to 0.02 resulting in an average error rate (on target and all non-target identities) of 0.04. This value is significant lower than the expected identification error rate IER (I_1).

We similarly proceed for the other detectors optimization (Figs. 6 and 7).

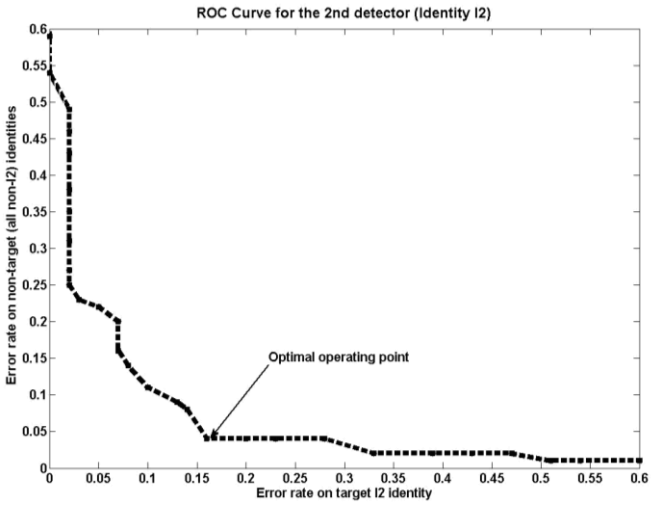

Fig. 6 Performance optimization for the 2nd detector (Identity I_2)

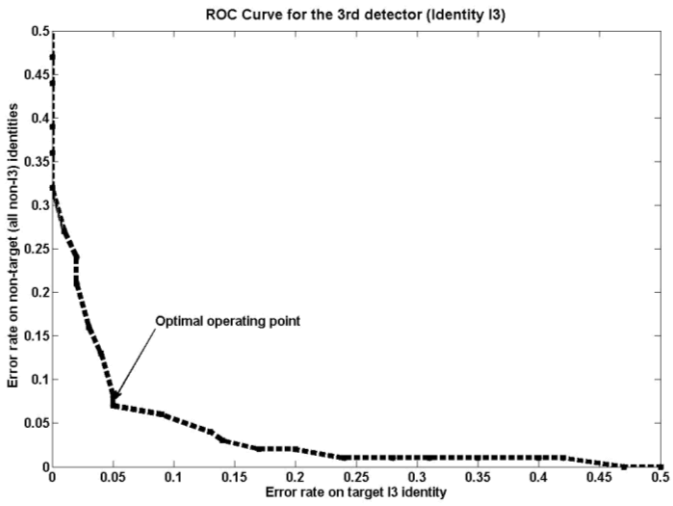

Fig. 7 Performance optimization for the 3rd detector (Identity I_3)

For the 2nd detector (which we trained for I_2 identity recognition), Fig. 6, we could see that its optimization is much more difficult. We achieved an average identification error rate IER (I_2) of 0.098, only for an increased threshold provided for identity I_2 (0.90 instead of the default threshold of 0.70). This value is only closest to the expected one for the default threshold. The cost is an increased identification error rate on I_2 class. However for security reason we have to limit the identification error rate on person I_2 to at most 0.20 and therefore a further optimization could not be realized only by this thresholding for person I_2 identification.

The operating points for the 3rd detector (identity I_3) are revealed on the ROC curve depicted in Fig. 7. In this case, the optimization is performed by reducing the threshold for I_3 detection from the default value of 0.70 to the best value which is 0.50. The average error rate IER (I_3) is 0.06 in this case, revealing again a significant improvement with respect to the initial expected value of 0.10 (found for the default threshold).

In all these 3 cases we performed the biometric detectors optimizations (for our identification designed system) only by thresholding on the classifiers outputs. However, the identification system is very sensitive not only to the fixed threshold, but also to the input biometric data quality, either for enrollment or for testing (authentication). Also we considered the classification rejection rate by fixing it to a fraction of 5% in the system design stage. This is important just to prevent or to minimize the identification errors due to external factors, which are not dependent on the classification algorithms accuracy. Finally the application security requirements are driving the whole optimization process in order to design and to implement a more efficient biometric-based security system.

VI. CONCLUSIONS

The biometric identification systems need to consider different optimization options depending on their applications requirements. The identification accuracy still remains an open issue as much as there are a lot of factors to be considered for an optimal solution design, providing a suitable trade-off between the provided performance improvement and the implementation costs. Also the feature number is an important issue to be considered.

Having in mind this challenges we proposed a method for biometric data classification in which the identification is relying on a special kind of classifiers called detectors. These detectors are only trained on target identities and this approach is most suitable for many biometric applications which have various security and / or cost / performance ratio-regarding requirements. If it is more important to precisely identify one person than all the other enrolled users of a critical resource (for instance a medical database), the detector-based approach seems to provide the best performances. On the other hand we applied our method on a significant reduced-sized feature space, although this feature space is a fused one, combining 2 different biometric features, fingerprint and palmprint respectively, within a simple pre-classification fusion rule. Therefore the performance improvements of our approach are resulting from the feature-level fusion and the thresholding detectors optimization.

Further research should be performed to include a more specialized pre-classification fusion rule, for instance by a careful exploring of feature correlation to get more performance improvements in identification.

REFERENCES

[1] Soviany S., Puscoci S., Jurian M., "A medical data biometric security model with multiple detectors", In Proceeding of the 32th National Conference on Medical Informatics ROMEDINF 2012, Timisoara, Romania, Nov. 15-17, 2012.

[2] Eleyan A., Demirel H., "Co-occurrence matrix and its statistical features as a new approach for face recognition", Turk J. Elec Eng. & Comp Sci, vol. 19, no. 1, 2011.

[3] Zucker S.W., Terzopoulos D., "Finding Structure in Co-Occurrence Matrices for Texture Analysis", Computer Graphics and Image Processing no. 12, 1980.

[4] Bino S., Unnikrishnan A., Kannan B., "Gray level Co-Occurrence Matrices: Generalization and some new features", International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), vol.2, no. 2, Apr. 2012.

[5] Jolliffe I. T., "Principal Component Analysis", Springer-Verlog, 1986.

[6] Theodoridis S., Koutroumbas K., "Pattern Recognition", 4th edition, Academic Press Elsevier, 2009.

[7] Devroye L, Gyorfy L., Lugosi G., "A Probabilistic Theory of Pattern Recognition", Springer, 1997.

[8] Jain A., Nandakumar K., Ross A., "Score Normalization in multimodal biometric systems", Pattern Recognition, The Journal of the Pattern Recognition Society, 38(2005).

[9] PerClass Training Course: Machine Learning for R & D Specialists, Delft, Netherlands.

[10] Duan K., Keerthi S., "Which Is the Best Multiclass SVM Method? An Empirical Study", Springer Verlag, Berlin 2005.