Fuziunea la nivel de caracteristici în sisteme biometrice optimizate pentru aplicaţii mobile

Feature-level Fusion in Biometric Systems optimized for Mobile Applications

Sorin Soviany[1], Virginia Săndulescu[1], Sorin Puscoci[1]

Rezumat: *Articolul abordează problema implementării fuziunii la nivel de caracteristici ca tehnologie multimodală în proiectarea de soluţii biometrice de securitate pentru aplicaţii mobile (de tip m-Banking, m-Commerce, m-Health sau m-Government). Se propune o metodologie de realizare a fuziunii la nivel de caracteristici în care procesul de combinare se realizează inter-modal (pentru tipuri biometrice diferite), dar în mai multe etape; metodologia de proiectare permite şi includerea unui stadiu suplimentar de fuziune intra-modală la nivel de caracteristici. Metodologia este aplicabilă pentru spaţii de caracteristici cu dimensionalitate redusă, cerinţă relevantă în cazul aplicaţiilor impunând constrângeri semnificative din punct de vedere al resurselor de procesare, stocare dar şi al vitezei de transmisie. Aceasta este cazul aplicaţiilor mobile.*

Cuvinte cheie: *fuziune de caracteristici, intra-modal, inter-modal, aplicaţii mobile*

Abstract: *The paper addresses the topic of feature-level fusion implementation as a multimodal technology for security biometric solutions design in case of mobile applications (such as m-Banking, m-Commerce, m-Health or m-Government). A feature-level fusion methodology is proposed with a combination process that is applied inter-modality (for different biometrics), but in several stages; this design framework allows to include an additional intra-modal feature-level fusion. The methodology is suitable for small-sized feature spaces; this is a relevant requirement for applications with significant constraints concerning processing and storing resources and transmission rate. This is the case of mobile applications.*

Keywords: *feature fusion, intra-modal, inter-modal, mobile applications*

1. Introduction

The data fusion process is the typical functionality for any multimodal biometric system. This operation provides a compact and meaningful representation of the similarity measures or feature sets achieved from several biometric traits in order to perform the individuals recognition. The resulted representation given by the multimodal fusion depends on the processing level in which it is performed. Actually there are 2 main categories of biometric fusion: *post-classification* and *pre-classification* fusion, respectively [1].

The *post-classification* fusion schemes are matching **score-level**, **rank-level** and **decision-level fusion** [1],[2],[3]. The combination is performed after the matching/classification process and operates on similarity scores (with various mathematical functional rules), identification degrees ranking or even on decisions (with voting schemes, AND or OR rules, fuzzy-based models) [1],[3]. This is the most implemented fusion in the actual multimodal biometric systems because of its low complexity [3],[4]. However, there are some drawbacks of this fusion,

[1]Institutul Naţional de Studii şi Cercetări pentru Comunicaţii, Bucureşti, România

concerning the achieved performance improvement that does not always meet all the security application requirements. In this case the multimodal biometric fusion is applied within an advanced stage of biometric data processing and this does not allow to efficiently exploit the most informative properties insight of the raw biometric data [4].

The *pre-classification* fusion schemes are **sensor-level** and **feature-level fusion** [1],[2]. The combination is performed before the matching stage of the human recognition and therefore it exploits the most informative properties of the raw data, either as individual samples from the biometric sensors, before any feature extraction/computing process (for **sensor-level fusion**) or as extracted features with certain algorithms for feature computing (for **feature-level fusion**). The **sensor-level fusion** is a more theoretically one, because its real application is strongly limited by the incompatibility of the sensors and raw samples [2],[3]. The **feature-level fusion** is more interesting than all the post-classification fusion rules because the fusion process combines the features sets that are extracted from several independent biometric data sources and this provides more discriminant information; this is a reliable reason to assume that the **feature-level biometric fusion** should ensure a higher performance improvement for the recognition process than the *post-classification fusion*. However, there are significant issues that limit the **feature-level fusion** implementation in real applications, despite of its promising potential performance improvements. This challenges are the following [1],[2],[4]:

- to find relationships among the various feature spaces for the biometric traits to be integrated. The task of providing a common representation of all the extracted features with an explicite insight of all possible correlations among the features is still limited by the complexity of the various feature extraction algorithms;
- the unavailability of the feature vectors structure for the commercial proprietary biometric systems. There are security reasons for this, however it is an issue that prevents an application developer or integrator to implement a reliable feature-level fusion rule;
- the dimensionality increasing, expecially when the fusion is based on the feature vectors concatenation. This leads to the curse of dimensionality and peaking with negative impact on the classification accuracy as concerning the classifier complexity and training size vs. feature space size ratio [4],[5]. This will reduce the security performances for the real application;
- the complexity of various feature extraction/computing techniques, with negative impact as concerning the incompatibility issues for the resulted features. Even the concatenation-based feature fusion is not always feasible for features that are not compatible, althought this feature-level fusion is not constrained by the same dimension of the input feature vectors:
- the homegeneity degree of the extracted/computed features for each of the integrated biometrics. This concerns either the feature vectors sizes and the numeric range of the features values. The various feature extraction/computing methods provide feature sets with values that are within different ranges and therefore not suitable for the fusion process.

Despite of the technical issues related to the **feature-level fusion** implementation for real multimodal biometric applications, some particular solutions are feasible to design high-performance multimodal biometric systems meeting constraints specific for applications. For instance, in the case of mobile applications such as m-Banking or m-Health, it is possible to design and implement feature-level fusion allowing to provide an optimal adjustment of the performance vs. complexity ratio. These approaches should consider the typical resources constrains of the mobile applications: storing, processing and transmission speed, respectively. This is especially true when the used terminals are smartphones, despite of the advanced in their operating systems. For mobile applications with smartphones, the typical biometrics that can be used for the end-users authentication are fingerprint, iris, face and voice. During the last few years the fingerprint-based authentication was introduced for several smartphones (Samsung Galaxy, i-Phones), although some security issues were already noticed concerning the fraudulent exploiting of the fingerprint patterns in order to get access to confidential data and other mobile services-related resources. These security issues justify the actual efforts to design and develop advanced solutions based on multimodal biometrics but customized for the mobile applications specific requirements.

In this paper we define a methodological framework suitable to design and implement a feature-level fusion for 2 biometric traits (in this case, fingerprint and iris). The combination process is applied inter-modality (for different biometrics), but in several stages; this design allows to include an additional intra-modal feature-level fusion. The methodology is suitable for small-sized feature spaces, and therefore for low computational complexity applications such as the mobile ones. The proposed methodology involves homogeneous feature vectors and this is why the fusion could be performed with a functional combination of the input vectors. However the concatenation-based approach for feature-level fusion could be considered within this design. The reduced feature space size allows to avoid the problems given by curse of dimensionality and peaking. Within this framework some feature space transformations are also considered in order to optimize the dimensionality to get resulting uncorrelated features and also to ensure the best class-separation within the further matching operation.

The remainder of this paper has the following structure: *Section 2* presents some recent related works concerning the biometric feature-level fusion; *Section 3* presents the proposed methodology for feature-level fusion design and implementation; *Section 4* concludes about the opportunities and challenges related to this biometric fusion, the advantages of the proposed methodology and its applicability for real cases, with focus un mobile applications.

2. Basic Developments and Related Works in Feature-level Biometric Fusion

The feature-level fusion is the process of biometric data combining in which the feature sets that are extracted from different primary biometric samples (raw data) are fused to generate a single feature vector. The resulted feature vector contains the most relevant discriminant features provided either:

- from *various human traits*: this is the typical **inter-modal feature-level fusion**, in which different feature sets derived from several biometric modalities are combined, or
- from *the same human trait* but with various sampling or feature extraction algorithms, eventually different parameterizations of the same feature extractor: this approach is the **intra-modal feature-level fusion**, in which the fusion is applied for the features that are extracted from the same biometric.

This fusion was mostly approached within some academic researches for various biometric traints and considering the suitable feature extraction methods [4]. However, it was not so much considered for real multimodal biometric applications development and implementation, and the reasons were given by Jain and Ross in [1]: the relationships among the various feature spaces, the lack of availability of the feature vectors in the actual commercial biometric systems and the curse of dimensionality problem that appears if the feature-level fusion is based on the input vectors concatenation.

However, the opportunity to exploit the most informative properties of the biometric samples within an earlier processing stage, just after the feature extraction but before the matching/classification stage provided a good reason to approach this fusion in order to find out suitable solutions for specific applications requirements.

The actual developments in feature-level fusion are based on the following approaches for the feature sets combination [1],[6]:

- *feature vectors concatenation* (figure 1): this design approach is applied when the input feature vectors (x,y) are ***not homogeneous*** and have ***different sizes*** (d_x, d_y). The biometric features ***heterogeneousness*** is provided by the different pre-processing methods for feature extraction. Despite of its easy implementation, the main drawback of this fusion is the dimensionality increasing for the resulted vector (z). The involved costs concerns the

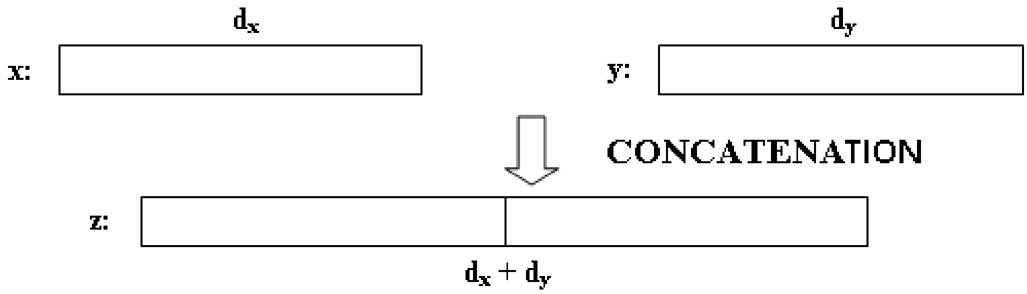

Figure 1: Concatenation-based Feature-level Fusion [3],[6]

computational complexity of biometric data processing and also the curse of dimensionality problem [5]. Another unfeasibility issue relates to the incompatibility of some extracted features. For example, in a fingerprint+face multimodal biometric system, the minutiaebased features (from fingerprint) and eigenface coefficients (from face) are incompatible [6],[7]. This approach is typical for the inter-modal feature-level fusion;

- *feature vectors functional combination* (figure 2): this design approach mandatorily requires the feature vectors ***homogeneity*** and also ***the same dimensionality*** of the feature spaces. The ***homogeneous*** feature vectors result from multiple samples acquisition for the same biometric, and/or with the same biometric data pre-processing algorithm for feature extraction/computing [1],[8]. Sometimes the feature extractor has different parameterizations. The most common functional fusion is the weighted averaging of the

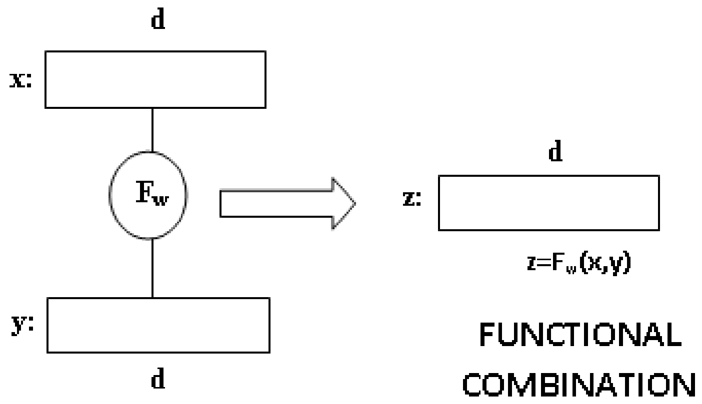

Figure 2: Functional combination-based Feature-level Fusion [3],[6]

biometric feature vectors components. This approach is typical for the intra-modal feature-level fusion but it could also be applied for the inter-modal feature-level fusion if the feature extraction algorithm is basically the same for several different biometrics and the only differences are given by the parameterization. However, the functional feature-level fusion is more constrained than the concatenation feature-level fusion.

On the other hand, Zhang et al. showed that the fusion-level fusion could be performed in the following ways [2],[4]:

- *the separate feature extraction for each biometric, followed by the true fusion of the resulted feature vectors* (figure 3[2]). In this approach the resulted features that are extracted from each biometric are directly fused using the concatenation or the functional combination, depending on the feature vectors homogeneity degree and also on the compatibility issues of the features [4];

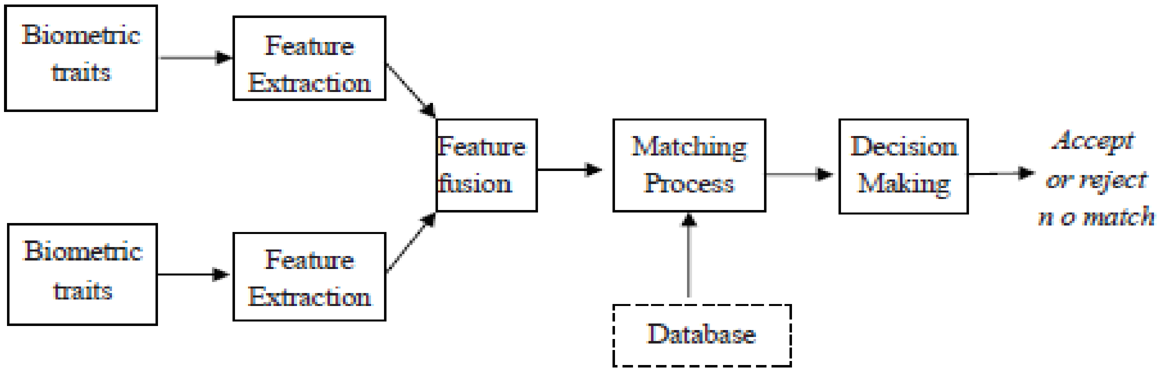

Figure 3: Feature-level Fusion with separate feature extraction algorithms for each biometric [2]

- *the fusion of the original images (biometric traits) followed by a single feature extraction algorithm application on the combined biometric trait* (figure 4 [2]). In this approach the overall feature vector containing the information from the original fused traits is a result of the single feature extraction procedure [4].

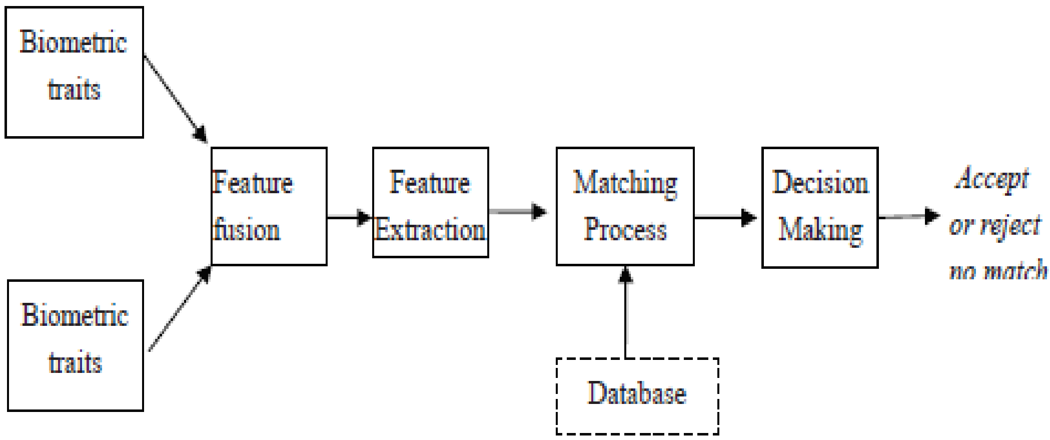

Figure 4: Feature-level Fusion with combined biometric traits (original images) and a single feature extraction algorithm [2]

A few examples of multimodal biometric systems with feature-level fusion are given as following.

In [9] an example of **intra-modal feature-level fusion** for multiple fingerprints is proposed. The feature-level fusion is applied for multiple instances of the same biometric trait (fingerprint), therefore it is an intra-modal fusion. The resulting feature vector is obtained by augmenting the input feature vectors and performing feature selection on the concatenated vector. This is an example of *concatenation-based feature-level fusion*.

An **inter-modal feature-level fusion** was developed and presented in [10] for fingerprint and iris. The unimodal and homogeneous features from these biometrics are combined using a Mahalanobis distance-based technique. Both the input feature vectors have the same size and the fused vector has the same length as the unimodal vectors, to avoid the high dimensionality problem. In this approaches 2 difference vectors are computed between each input biometric vector and its nearest match with the corresponding reference template, using the Mahalanobis distance. After a normalization step, the final fused vector is obtained as the mean value of these 2 difference vectors. The fused fingerprint+iris vector is then used for training and classification with SVM (Support Vector Machine) classifier. This is an example of a *functional combination-based feature-level fusion*.

Another example of an **inter-modal feature-level fusion** is given in [11]. In this case the 2 biometrics are face and palmprint. The feature extractors for both biometrics are based on PCA (Principal Component Analysis) and CCA (Canonical Correlation Analysis) and used to carry out feature fusion and to get correlation information. This is a *functional-combination* of the input feature vectors but with a special focus on their components correlation.

A *concatenation-based* **inter-modal feature-level fusion** algorithm is presented in [12]. In this case the input feature vectors are provided from face and hand biometrics. For these heterogeneous feature vectors the most obvious approach of the feature fusion is to perform their concatenation together with the normalization of the input features. As much as the concatenation-based fusion has the typical drawback of a very large dimensional resulting feature space, the explicit feature selection is required in this approach in order to reduce the dimensionality of the fused feature vector.

The previous works are only a few of some recent researches in the area of the feature-level biometric fusion as basis to design and develop high-performance biometric systems for real applications with various constraints. Despite of them, the pre-classification fusion (and especially the feature-level fusion) still remains a hot R&D topic for multimodal biometric systems design, and this is because of its challenges. On the other hand, the promising potential of a significant biometric recognition performance improvement comparing with the post-classification fusion rules is a reliable reason for the actual efforts concerning this design option concerning the multimodal biometric systems development.

3. The Feature-level Intra- and Inter-Modal Biometric Fusion Methodology

Starting with the previous analysis of the feature-level biometric fusion issues, we define a methodological framework to perform this biometric data fusion while considering both main approach, concatenation and functional combination, respectively, with intra-modal and inter-modal application. The methodology specification for the feature-level fusion requires the following elements:

- *the multimodal biometric system together with its main functional components specification that are required for the* ***feature-level fusion***;
- *the feature generation process*, meaning all the required operations that provide the extracted features from the input biometric traits;
- *the* ***feature-level fusion*** *procedure* with several operational cases handling.

3.1 The Multimodal System Architecture

Let's consider the multimodal biometric system given in [13] as a reliable support of identification methods for mobile application based on smartphones. The system architecture (figure 5) includes 3 biometric functional components: fingerprint, iris and voice recognition,

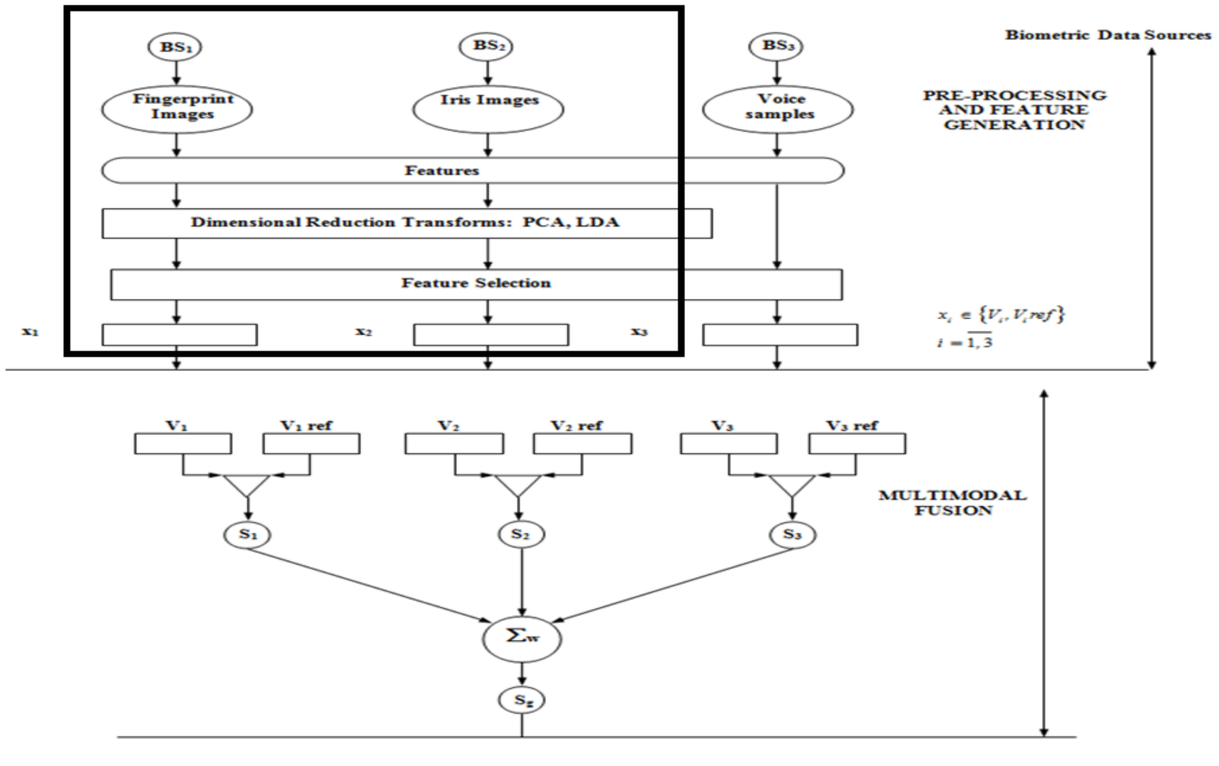

Figure 5: The multimodal biometric system architecture for mobile applications [13]

respectively [13]. This is the full system architecture with a final post-classification (matching score-level) fusion of all the 3 biometrics. For the feature-level fusion we only consider the first 2 biometrics, fingerprint and iris. This is because the applied feature extraction algorithms are quite similar in our approch, as much as the proposed methodology focuses on small-sized festure spaces that should be suitable for the mobile applications specific constraints. The feature homogeneity could be only provided for fingerprint and iris while applying textural approaches for statistical features computing that are typically for image processing.

We consider the fingerprint and iris data pre-processing and processing operations in order to generate the input feature vectors for the **feature-level fusion**. The third biometric will be separatelly processed and it will be subject to the final score-level fusion. Therefore the functional components are the following:

- ***feature generation*** for both biometrics;
- ***feature-level fusion*** with several operational cases.

3.2 The Feature Generation

The feature generation is the activity in which the biometric raw data (primary samples) are processed to perform the following tasks:

- *Filtering and noise reduction* from the primary biometric samples (*pre-processing stage*);
- *Regions-of-Interest (ROIs) selection* from the original image;
- *The features computation*;
- *The features normalization*;
- *The feature space transformation* for dimensionality reduction and/or features uncorrelation and/or maximization of the class-discriminant property of the extracted features;
- *The feature selection* in order to only keep the most informative features for the futher fusion process.

For the scope of the designed methodological framework we only consider the following basic tasks belonging to the feature generation stage, according to figure 6:

- *T1*: ***a fixed number of ROIs selection*** within each of the input images;
- *T2*: ***the 2nd order statistical computing*** from each of the previously selected ROI. This task also includes the *normalization* of the extracted features, operation that ensures the common values ranges and scaling;
- *T3*: ***the dimensionality reduction*** with the following sub-steps:
  - *a secquence of feature spaces tranformations* providing the best features with low correlations, high variance, high degree of statistical independence and/or high class separability;
  - *an explicite feature selection procedure* that provides the optimal feature subsets from both biometrics (fingerprint and iris) with the highest discriminant performance.

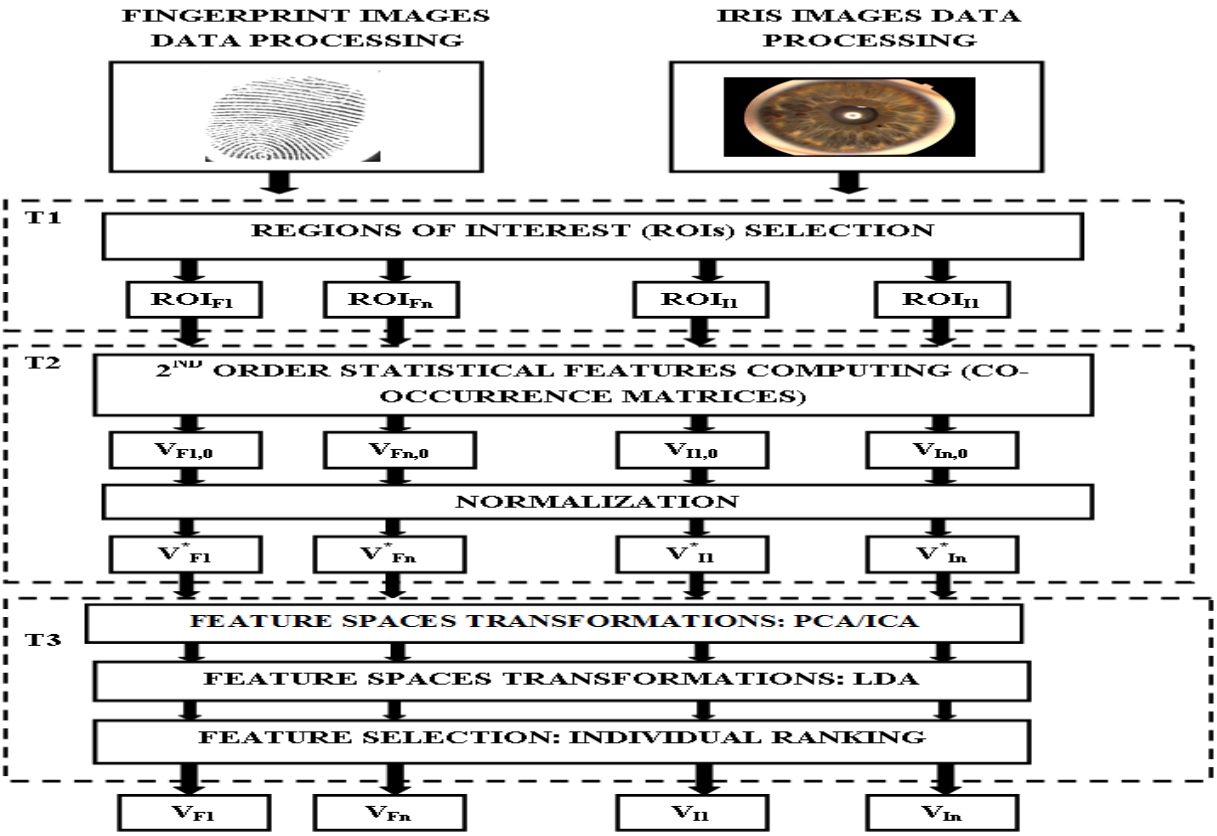

Figure 6: The Feature Generation process

3.2.1 Regions of Interest Selection

The first operation within the whole feature generation process is the manual selection of the ROIs (regions of interest) within each of the original input images (task T1).

The **Region of Interest** (*ROI*) is a certain zone within the input image that contains the most useful information for the application. In this case, the ROIs within the iris and fingerprint images contain the most discriminants details that allow to perform an accurate recognition of the individuals by optimally exploiting the informative properties of the extracted features. Actually for the biometric applications the selected ROIs within the input images (fingerprint, iris and any other biometric requiring image processing) are used to extract just the most informative features reducing the computational complexity of the further processing stages. This is especially important for applications with severe constraints concerning storing, processing and transmission speed, as the case of mobile applications. Therefore the regional approach for the further feature extraction is useful to ensure a reduced feature space just before any explicit feature selection.

For both biometrics that are considered to perform the feature-level fusion we apply a *manual procedure* to define and select the useful ROIs within the input fingerprint and iris images, respectively. In a manual procedure for ROI selection the application designer/developer explicitely defines the most useful region to be cropped from the original image by specifying its coordinates. This approach is different from the automatic methods in which one or several pre-defined regions are used for feature extraction, without the biometric application end-user or designer/developer involvement.

The *manual procedure of ROI definition/selection* requires the following operations:

- **fixing the number of ROIs for selection within the original image**. In our design we use the same number of ROIs for both biometrics (fingerprint, iris): $n_{ROI,F} = n_{ROI,I} = n = 2$;
- **defining the initial coordinates and their offsets for the rectangular areas** that corresponds to the meaningful regions within the original input images: $x_k, x_k + \Delta x_k, y_k, y_k + \Delta y_k, k \in \{F, I\}$, where x_k, y_k are the original coordinates of the region and Δx_k, Δy_k are their corresponding offsets; F is the label for the fingerprint source, and I is the label for the iris source. An alternative definition of the ROI coordinates is based on the direct specification of the upper-left and lower-right points, without offsets. These meaningul regions provide the most useful information for the biometric application;
- **defining and applying the mask matrices** allowing to extract the previously selected meaningful regions from the original input image. The mask matrix is given by [14]

$$M_k(i_k, j_k) = \begin{cases} 1, for\ x_k \le i_k \le x_k + \Delta x_k, y_k \le j_k \le y_k + \Delta y_k \\ 0,\ \text{otherwise} \end{cases}, k \in \{F, I\} \quad (1)$$

  where x_k, $x_k + \Delta x_k$, y_k, $y_k + \Delta y_k$ are the offset-based coordinates for the defined rectangular region of interest (ROI). The mask matrix size is the same as for the original image (given as a pixels array representation).

This regional approach for the feature generation allows to work with a small-dimensional feature space that is most suitable for the specific resource constraints of the mobile applications (data storing, processing and transmission speed) [4].

This manual ROI selection is applied on the gray-scale converted images. The original color images are transformed into the gray-scale images using the procedure given in [15]. This is suitable for the further textural analysis within the feature extraction process.

The results of this processing step are the extracted ROIs from each of the input images:

- $ROI_{F1},\dots,ROI_{Fn}$ for fingerprint;
- $ROI_{I1},\dots,ROI_{In}$ for iris.

3.2.2. Statistical Features Computing and their Normalization

From each of these extracted regions of interest the 2[nd] order statistical features are computed based on the co-occurence matrices (task T2). This operation provides n feature vectors for each of these 2 biometrics (fingerprint and iris, respectively).

These features describe the gray levels distribution within the input image [5], [16]; in our case, the input images are the previously extracted ROIs from the original fingerprint and iris images. Then the 2nd order statistical features are directly derived from the co-occurence matrices that are generated for each ROI [4]. These statistical features are pixel-pairwise computed [14].

A co-occurence matrix C is an array (or 2D histogram) in which each element evaluates the probability of a certain gray-level for one pixel in the input image (here ROI_{kj}, where $k \in \{F, I\}$ and $j = \overline{1,n}$, n is the number of the regions of interest that were previously selected and extracted during the task T1), while another displaced pixel exhibit another gray-level [14]. The co-occurrence matrix mathematical definition is the following [17], [18], [19]:

$$C_{\Delta x_{kj},\Delta y_{kj}}(m_{kj}, n_{kj}) = P\left\{ROI_{kj}(x_{kj}, y_{kj}) = m_{kj}, ROI_{kj}(x_{kj} + \Delta x_{kj}, y_{kj} + \Delta y_{kj}) = n_{kj}\right\}, k \in \{F, I\}, j = \overline{1,n} \quad (1)$$

This definition is applied considering as input images the selected ROIs from fingerprint and iris original samples. Here Δx_{kj} and Δy_{kj} are the pixels horizontal and vertical displacements, respectively. k is the index of the biometric trait (fingerprint, iris) and n is the index of the selected ROI from the original image.

The co-occurence matrices allow to efficiently exploit the underlying texture of the original image for the further feature extraction [16], [18]. The resulting co-occurence matrices-based feature extractor has the following parameterization [16], [19]:

- *GLB* (Gray-Level Bins number) is a parameter that could be adjusted to provide many significant values and less null values of the co-occurrence matrix elements, in order to ensure the most informative textural features from the input image (ROI). For our design the following setting is considered: $GLB_{Fj} = 4, j = \overline{1,n}$ for fingerprint image ROIs and $GLB_{Ij} = 6, j = \overline{1,n}$ for iris image ROIs. Further adjustments of this parameter will allow to significantly modify the resulted features number, ensuring an earlier dimensionality reduction, according to the specific application requirements;
- *OFFS* is an offset parameter that evaluates the displacement distance, actually the number of pixels between the pixels pairs that are used to generate the co-occurence matrix. It is important for this amount to not exceed a certain thresholding value; this is becase if this spacing is too high the resulting overall number of pixel pairs becomes too small and the achieved useful information is poor. For our design the suitable displacement threshold is $OFFS_{kj} = 2, k \in \{F, I\}, j = \overline{1,n}$.

These 2 parameters allows together to optimize the feature spaces dimensionalities in order to be suitable for the mobile applications specific constraints. On the other hand, the feature extraction algorithm is essentially the same for both integrated biometrics (fingerprint and iris). This provides a certain homogeneity degree of the resulted features and allows to apply a functional combination-based fusion, not only the typical concatenation-based feature-level fusion [4].

So far the resulted feature vectors are the following: $V_{F1,0},...,V_{Fn,0}$ for fingerprint and $V_{I1,0},...,V_{In,0}$ for iris. Their sizes derive from the parameterization of the co-occurrence matrices-based feature extractors, as following, for fingerprint and iris:

$$l_{Fj,0} = size(V_{Fj,0}) = \left(GLB_{Fj}\right)^2, j = \overline{1,n} \quad (2)$$

$$l_{Ij,0} = size(V_{Ij,0}) = \left(GLB_{Ij}\right)^2, j = \overline{1,n} \quad (3)$$

For both biometrics the feature extractor parameterization and design provides the same features number from each ROI. This condition is mandatory for the functional combination of the resulted vectors. Otherwise, if each ROI generates feature vectors with different sizes, only the concatenation-based intra-modal feature-level fusion is feasible.

The next operation is the feature normalization providing the homogeneity of the numerical ranges of the features. The normalization model is based on the simple sigmoid function providing output values within the range [0,1]:

$$V_{kj}^{*} = f(V_{kj,0}) = \frac{1}{1+\exp(-V_{kj,0})}, j = \overline{1,n}, k \in \{F, I\} \quad (4)$$

where $V_{kj,0}$ are the feature vectors generated from the co-occurrence matrices and V_{kj}^{*} are the normalized feature vectors. Their sizes are the same:

$$l_{Fj,0} = size(V_{Fj,0}) = size(V_{Fj}^{*}) = l_{Fj}^{*}, j = \overline{1,n} \quad (5)$$

$$l_{Ij,0} = size(V_{Ij,0}) = size(V_{Ij}^{*}) = l_{Ij}^{*}, j = \overline{1,n} \quad (6)$$

After completing the task T2 the resulted feature spaces dimensionalities are the following:

- for *fingerprint*, 2 feature vectors, each of them containing 16 elements (16 features that are extracted from each ROI): $l_{Fj}^{*} = l_{Fj,0} = 16, j = \overline{1,2}$;
- for *iris*, 2 feature vectors, each of them containing 36 elements (36 features that are extracted from each ROI): $l_{Ij}^{*} = l_{Ij,0} = 36, j = \overline{1,2}$

3.2.3 Feature Space Transformation for Dimensionality Optimized Reduction and Feature Selection

The feature spaces optimization is the process of dimensionality adjustment in order to customize the feature space according to the real application requirements and constraints. Actually an initial adjustment of the feature space dimensionality is already done even during the feature extraction/computing process (task T2), through the suitable choice of GLB/OFFS ratio. During the task T3 the feature space optimization is performed with the following operations sequence:

- ***unsupervised feature space transformation*** with *PCA* (*Principal Component Analysis*) or *ICA* (*Independent Component Analysis*) in order to provide the features with the highest variance degree, the uncorrelated features or even the statistical independent features (if applying *ICA*). *PCA* is an unsupervized algorithm that sometimes does not provide the optimal features for the class separability. To prevent this situation, for the proposed methodological framework *PCA* is applyed on the weighted covariance matrix S that is computed for C classes according to [4]:

$$S = \sum_{i=1}^{C} f_i \cdot S_i \tag{7}$$

in which: S_i is the class covariance matrix for the data belonging to the class i; f_i is the relative frequency (or prior probability) for the biometric class i within the overall training set. The number of classes is depending on the biometric application:
  - ➢ *C=2* (binary classification): the case of *verification process* or *target-vs.-non-target classification* of biometric data (a particular identification process with focus on the most important user recognition, while all the other enrolled users represent the non-target class);
  - ➢ *C>2* (multi-class classification): the case of *identification* process (in which the designed system should guess the true identity of a real person, only based on his/her biometric credential, without any username or other identifier). The overall classes number includes the identities of the enrolled users together with the unknown class that corresponds to an individual that was not previously registered within the system database;
- ***supervised feature space transformation*** with *LDA* (*Linear Discriminant Analysis*) that only retains the best features providing the optimal class discrimination. This is a linear transformation w that maximize the Fisher criterion, actually the inter-class vs. intra-class variance *J(w)* given by[4], [5]:

$$J(w) = \frac{\sigma^2_{inter-class}}{\sigma^2_{intra-class}} = \frac{w^T \cdot S_B \cdot w}{w^T \cdot S_W \cdot w} \tag{8}$$

where S_B is the between-class scatter matrix and S_W is the within-class scatter matrix;
- ***feature selection*** that provides the most relevant features from the original set, for outliers and redundancies removal. The further dimensionality reduction is ensured by retaining the most informative features. For this design the following non-exhaustive and non-optimal feature selection strategies are considered: *forward-searching*, *backward-searching*, *floating-search*, *individual ranking* and *random feature selection* [4]. The *exhaustive search* for relevant features always provides the optimal feature subsets but with the cost of execution speed and response time. The evaluations for several datasets show that the optimal execution time is ensured by the individual ranking approach if the original feature space size does not exceed 50.

The resulted dimensionality after completing the task T3 is the following:

$$l = size(V_{Fj}) = size(V_{Ij}) = 10,\ j = \overline{1,2}$$

So for each biometric (fingerprint and iris) the methodological framework only retains 10 features; this is suitable, from the complexity point of view, with the resource-constrained requirements of the mobile applications (data storing, processing and transmission speed).

On the other hand, the same dimensionality of the feature vectors that are derived from each ROI for both biometrics allows to consider not only the concatenation, but also the functional combination of the feature vectors. This is a flexibility advantage of this methodological framework.

3.3 The Feature-level Fusion

For the feature-level fusion we consider the main approaches within the proposed methodological framework, also with several design options depending on the homogeneity degree of the generated feature vectors: inter-modal and intra-modal feature-level fusion, either separated and then integrated into an overal feature-level model:

- the *inter-modal feature-level fusion*, in which the only combination performs between the feature vectors that derive from the different biometrics;
- the *intra-modal feature-level fusion*, in which the combination process considers the feature vectors that derive from the same biometric;
- the *full feature-level fusion* integrating the previous approaches: inter-modal and intra-modal fusion.

3.3.1 The Inter-Modal Feature-level Fusion

The inter-modal feature-level fusion is the biometric data fusion process that combines the feature vectors provided by several different biometric traits.

For this design a *multi-stage* inter-modal fusion process is considered, actually an inter-modal feature-level fusion with 2 operational stages, according to the model depicted in figure 7. The 2 stages of the inter-modal fusion are the following:

- ***the local inter-modal fusion***, in which the fusion process combines each ROI-extracted feature vector from fingerprint and iris. The only condition is to have the same number of ROIs for both biometrics;
- ***the global inter-modal fusion***, in which the fusion process combines the feature vectors that results from the local inter-modal fusion.

Both of these inter-modal feature-level fusion stages could be approached with concatenation or functional combination, depending on the homogenity and compatibility of the extracted features.

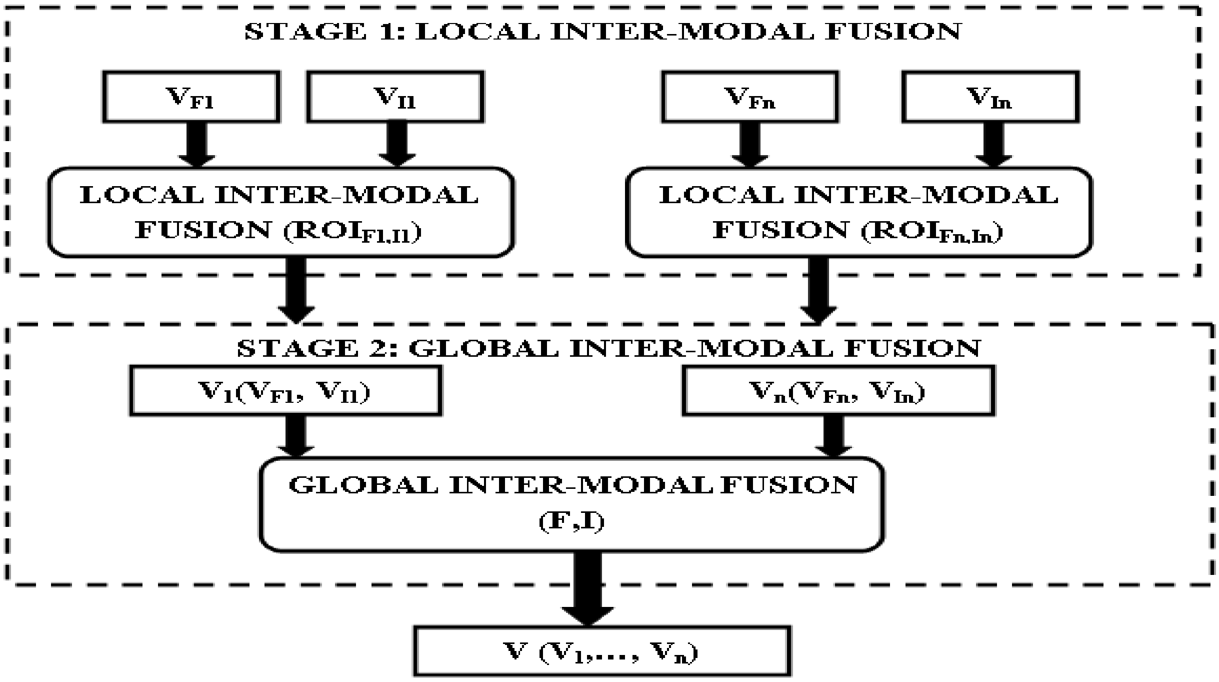

Figure 7: The multi-stage inter-modal feature-level fusion

A) The local inter-modal fusion

In this case, the input feature vectors pairs are the following, according to figure 7:

- for ROI_{F1}, ROI_{I1}: V_{F1}, V_{I1};
- for ROI_{F2}, ROI_{I2}: V_{F2}, V_{I2};
- for ROI_{Fn}, ROI_{In}: V_{Fn}, V_{In}

The local inter-modal fusion combines the vectors from each of the *n* pairs, generating *n* feature vectors, one fused feature vector for each ROI within fingerprint and iris, respectively. The resulted feature vectors are the following: $V_1(V_{F1}, V_{I1}),\dots, V_n(V_{Fn}, V_{In})$.

Depending on the dimensionality and homogeneity degree, the feature fusion is performed in one of the following ways:

- **the functional combination of the input vectors** within each of the *n* pairs:

$$V_j = f(V_{Fj}, V_{Ij}; w_F, w_I), j = \overline{1,n} \qquad (9)$$

in which V_j is the fused feature vector (fingerprint+iris), f is the selected functional combination model and w_F, w_I are the assigned weights for the fused biometrics. The main condition is to have the same dimensionality for both biometric feature spaces. The typical functional combinations of the feature vectors are the weighted sum and average:

$$V_j = w_F \cdot V_{Fj} + w_I \cdot V_{Ij}, j = \overline{1,n} \tag{10}$$

$$V_j = \frac{w_F \cdot V_{Fj} + w_I \cdot V_{Ij}}{w_F + w_I}, j = \overline{1,n} \tag{11}$$

Other functional models could be applied depending on the experimental data and the application requirements;

- **the concatenation of the input vectors** for each of the n pairs:

$$V_j = [V_{Fj} \mid V_{Ij}], j = \overline{1,n} \tag{12}$$

This is the most general inter-modal feature-level fusion because it is not constrained by the dimensionality of the input feature vectors; their sizes could be different. The main drawback is the increased size of the resulted fused vector, requiring sometimes an additional feature reduction step, depending on the applications requirements concerning the computational complexity. Also the concatenation of the input vectors requires some care about the incompatibility among the fused features.

B) The global inter-modal fusion

The input for the 2nd stage of the inter-modal feature level fusion is given by the vectors V_j that result from the local inter-modal fusion stage: $V_j\left(V_{Fj}, V_{Ij}\right), j = \overline{1,n}$. The feature-level fusion is again performed through the functional combination or concatenation:

- **the functional combination-based fusion**: is a reliable design option when the input feature vectors have the same size and exhibit a high homogeneity degree (provided, for example, by the same feature extraction algorithm). The functional fusion rule is

$$V = F(V_1, ..., V_n) \tag{13}$$

where V is the final fused vector and F is the functional model for the input feature vectors combination. The most commonly applied functional model is a weighted sum:

$$F(V_1, ..., V_n) = \sum_{j=1}^{n} F_j(V_j; w_j) \tag{14}$$

In this case, the sum's terms are the weighted functional models F_j for the input feature vectors V_j. The weights w_j are applied for each ROI extracted and fused feature set;

- **the concatenation-based fusion**: is the most applied design option for the feature-level fusion:

$$V = [V_1 \mid V_2 \mid ... \mid V_n] \tag{15}$$

Typically it requires an additional feature selection in order to provide a suitable dimensionality of the resulted feature space. This is required in order to avoid the curse of dimensionality and peaking issues [5].

3.3.2 The Intra-Modal Feature-level Fusion. The Full Intra- and Inter-Modal Feature-level Fusion

Another approach for feature-level fusion is to combine the several feature sets from the same biometric; this is **the intra-modal feature-level fusion**. This procedure should be actually integrated into a **full intra- and inter-model feature-level fusion** with 2 operational stages like in the previous **inter-modal model**. The difference is that in the 1st stage the combination process applies on several feature sets from the same biometric (fingerprint and iris, respectively), while in the previous case the 1st stage fusion performs on the ROI-based feature vector, actually taking the feature sets that are extracted from each region of interest and inter-modally combining them between the 2 biometrics.

The full model is depicted in figure 8.

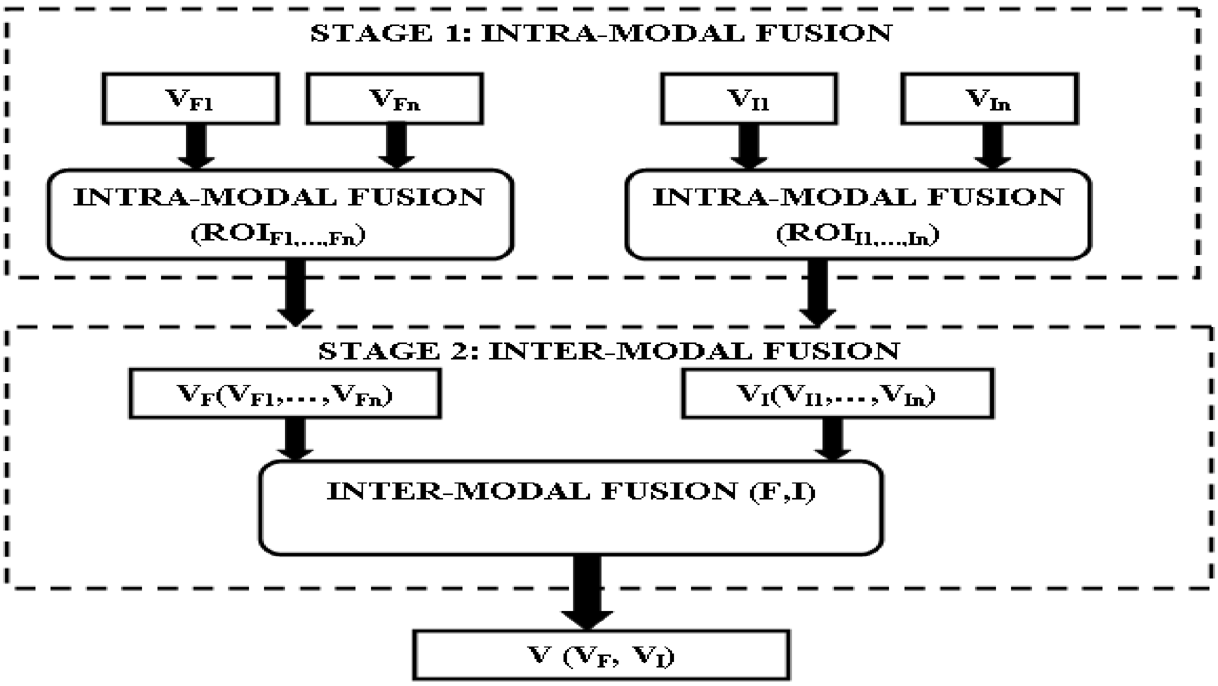

Figure 8: The multi-stage intra- and inter-modal feature-level fusion

The 2 stages are the following:

- ***the intra-modal fusion***, in which several feature vectors derived from the same biometric are fused in order to get a single representation. Each of the fused feature vectors is derived from a certain ROI from the same biometric input;

- ***the inter-modal fusion***, in which the fusion is applied on the feature vectors that result from the intra-modal fusion stage.

A) The intra-modal fusion

The input feature vectors that are subject of this fusion process are the following, according to figure 8:

- for biometric 1 (fingerprint, F): V_{F1}, V_{F2},..., V_{Fn};
- for biometric 2 (iris, I): V_{I1}, V_{I2},..., V_{In}

The intra-modal fusion combines the *n* feature vectors for each biometric, generating one vector for fingerprint and one vector for iris, respectively. The resulted feature vectors are the following: $V_F(V_{F1},\dots,V_{Fn})$ and $V_I(V_{I1},\dots,V_{In})$.

Again, the fusion is feasible in the following ways, depending on the equal dimensionalities and features homogeneity and compatibility degrees:

- **the functional combination of the input vectors**, with the general model given by
$$V_k = f\left(V_{k1},\dots,V_{kn}\right), k \in \{F, I\} \qquad (16)$$
This is only feasible if all the *n* input feature vectors have the same size. The typical rules are the weighted sum or average, but also other functions could be considered. For this design option the weighting is focused on the ROI as direct source of the feature vector for the intra-modal fusion process. So the ROI selection is essential to provide the best features and also to ensure an accurate intra-modal feature-level process. The resulted models are:
$$V_k = \frac{\sum_{j=1}^{n} F_{kj}(V_{kj}; w_{kj})}{\sum_{j=1}^{n} w_{kj}}, k \in \{F, I\} \qquad (17)$$
- **the concatenation of the input vectors** with an inherent dimensionality increasing:
$$V_k = [V_{k1} \mid V_{k2} \mid \dots \mid V_{kn}], k \in \{F, I\} \qquad (18)$$
A further feature selection or other transformations for dimensionality reduction is usually required in this case.

B) The inter-modal fusion

In this stage the 2 feature vectors that result from the previous stage, V_F and V_I, are combined to get a final fused vector containing the most relevant discriminant features from both biometrics. The combination process could be performed with the same 2 approaches:

- **the functional combination**, requiring that the input feature vectors should meet the same conditions concerning the features homogeneity and the same dimensionality. The general fusion model is given by

$$V = F(V_F, V_I) \qquad (19)$$

As in the previous sub-cases, the particular models are based on sum or averaging, sometimes with a certain weighting, depending on the real application:

$$V = \frac{w_F \cdot V_F + w_I \cdot V_I}{w_F + w_I} \qquad (20)$$

- **the concatenation** of the 2 input feature vectors:

$$V = [V_F \mid V_I] \qquad (21)$$

The same remarks should be made here as concerning the resulted feature space dimensionality increasing that requires a further adjustment through feature selection or other transformation able to only retain the most discriminant features and to meet the application specific constraints.

4. Conclusions

In this paper a methodological framework is defined to design and implement the feature-level fusion for biometric authentication mechanisms that are particularly focused on mobile applications. The mobile applications such as m-Banking and m-Health require an optimal design for security but also with a suitable computational complexity in biometric data processing. This is because of the resource-related constraints of the mobile terminals (data storing, processing and the transmission speed).

On the other hand, the feature-level fusion still remains a big challenge for the actual multimodal biometric systems design. This is because of the various feature extraction methods and the unavailability of the feature vector structure in case of the commercial biometric devices and applications. However, the feature-level fusion provides a significant potential for further performance improvements in biometric recognition because the combination process perfoms just in an earlier stage of data processing, after the feature extraction but before the matching/classification and recognition decisions tasks.

The proposed framework includes several design options for the feature-level fusion, with differentiated cases. The methodology considers the 2 basic feature-level approaches (feature vectors functional combination and concatenation, respectively) . Also the whole fusion process is applied in a hierarchical way, with several stages of fusion.

An important issue with this methodological framework is that it operates on reduced feature spaces dimensionalities. This design option is useful to avoid the curse of dimensionality and peaking for the biometric data classification.

Further issues regarding the feature-level fusion relate to the homogeneity degree of the extracted features and their compatibility. These issues should be carefully considered for an optimal design and implementation according to the particular applications requirements.

References

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

[2] D. Zhang, F. Song, Y. Xu, Z. Liang: *Advanced Pattern Recognition Technologies with Applications to Biometrics*, Medical Information Science Reference, IGI Global, 2009

[3] S. Soviany: Ph.D. Thesis *Decision Optimization in Biometric Identification Systems*, University of Pitesti, January 2013

[4] S.Soviany, C.Soviany, S.Puşcoci: *A Multimodal Biometric System with Several Degrees of Feature Fusion for Target Identities Recognition*, The 2016 International Conference on Security and Management (SAM'16), Las Vegas, USA, July 25 – 28, 2016. In Press (accepted for presentation and publication in Conference Proceeding)

[5] S.Theodoridis, K. Koutroumbas: *Pattern Recognition, 4th edition*, Academic Press Elsevier, 2009

[6] S.Soviany, S.Puşcoci: *The Multimodal Fusion in Security Systems Design*, Telecomunicaţii Nr. 1/2015

[7] A.Jain, A.Ross, S.Prabhakar: *An Introduction to Biometric Recognition*, IEEE Transaction on Circuits and Systems for Video Technology, Special Issue on Image- and Video-Based Biometrics, Vol. 14, Nr. 1, 2004

[8] R.Snelick, M.Indovina: *Multimodal biometrics: issues in design and testing*, Proceedings of 5th International Conference on Multimodal Interfaces, Canada, 2003, pp. 68-72

[9] A. J. Jacob, N. T. Bhuvan, S, M. Thampi: *Feature Level Fusion using Multiple Fingerprints*, IJCA Special Issue on “Computational Science - New Dimensions & Perspectives” NCCSE, 2011

[10] U. Gawande, M. Zaveri, A. Kapur: *A Novel Algorithm for Feature Level Fusion Using SVM Classifier for Multibiometrics-Based Person Identification*, Hindawi Publishing Corporation Applied Computational Intelligence and Soft Computing Volume 2013

[11] Z. Wang, C. Liu, T. Shi, Qun Ding: *Face-Palm Identication System on Feature Level Fusion based on CCA*, Journal of Information Hiding and Multimedia Signal Processing ,Ubiquitous International Volume 4, Number 4, October 2013

[12] S.K.Bhardwaj: *An Algorithm for Feature Level Fusion in Multimodal Biometric System*, International Journal of Advanced Research in Computer Engineering &

Technology (IJARCET) Volume 3 Issue 10, October 2014
[13] S. Soviany, V. Săndulescu, S. Puşcoci: *A Multimodal Biometric Identification Method for Mobile Applications Security*, ECAI 2016 - International Conference – 8th Edition Electronics, Computers and Artificial Intelligence, Ploiesti, România, 30 June -02 July, 2016
[14] S. Soviany, C. Soviany: *A Biometric Security Model with Identities Detection and Local Feature-level Fusion*, The 2013 International Conference on Security and Management (SAM'13), World Academy of Science, Las Vegas, SUA, July 22-25, 2013
[15] Bhattacharyya D.,Das P.,Bandyopadhyay S.K., Kim T.: *IRIS Texture Analysis and Feature Extraction for Biometric Pattern Recognition*, International Journal of Database Theory and Application, vol. 1, nr. 1, pp. 53-60, December 2008.
[16] S. Soviany, C. Soviany, S. Puşcoci: *An Optimized Iris Recognition System for Multi-level Security Applications*, The 2014 International Conference on Security and Management (SAM'14), World Academy of Science, Las Vegas, SUA, July 21-24, 2014.
[17] A.Eleyan, H.Demirel: *Co-occurrence matrix and its statistical features as a new approach for face recognition*, Turk J Elec Eng & Comp Sci, Vol.19, Nr.1, 2011
[18] S.W. Zucker, D. Terzopoulos: *Finding Structure in Co-Occurence Matrices for Texture Analysis*, Computer Graphics and Image Processing nr. 12, 1980
[19] S. V. Bino, A. Unnikrishnan, B.Kannan: *Gray level Co-Occurrence Matrices: Generalisation and some new features*, International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), Vol.2, No.2, April 2012