

# Class-Specific Discriminant Non-negative Matrix Factorization for Frontal Face Verification

Stefanos Zafeiriou, Anastasios Tefas, Ioan Buciu, and Ioannis Pitas

Aristotle University of Thessaloniki,  
Department of Informatics 54124  
Thessaloniki, Greece

{dralbert, tefas, nelu, pitas}@aiaa.csd.auth.gr

**Abstract.** In this paper, a supervised feature extraction method having both non-negative bases and weights is proposed. The idea is to extend the *Non-negative Matrix Factorization* (NMF) algorithm in order to extract features that enforce not only the spatial locality, but also the separability between classes in a discriminant manner. The proposed method incorporates discriminant constraints inside the NMF decomposition in a class specific manner. Thus, a decomposition of a face to its discriminant parts is obtained and new update rules for both the weights and the basis images are derived. The introduced methods have been applied to the problem of frontal face verification using the well known XM2VTS database. The proposed algorithm greatly enhance the performance of NMF for frontal face verification.

## 1 Introduction

Face recognition/verification has attracted the attention of researchers for more than two decades and is among the most popular research areas in the field of computer vision and pattern recognition. The most popular among the techniques used for frontal face recognition/verification are the subspace methods. The subspace algorithms consider the entire image as a feature vector and their aim is to find projections (bases) that optimize some criterion defined over the feature vectors that correspond to different classes. Then the original high dimensional image space is projected into a low dimensional one. The classification is usually performed according to a simple distance measure in the final multidimensional space.

Various criteria have been employed in order to find the bases of the low dimensional spaces. Some of them have been defined in order to find projections that best express the population (e.g. *Principal Component Analysis* (PCA) [1], NMF [2], *Local Non-negative Matrix Factorization* (LNMF) [3]) without using the information of how the data are separated to different classes. Another class of criteria is the one that deals directly with discrimination between classes (e.g. *Linear Discriminant Analysis* (LDA) [4]).

A subspace method that aims at finding a face representation by using basis images without using class information is NMF [2]. The NMF algorithm, like PCA, represents a face as a linear combination of bases. The difference with PCA is that it does not allow negative elements in both the basis vectors and the weights of the linear combination.

This constraint results to radically different bases than PCA. On one hand the bases of PCA are eigenfaces, some of which resemble distorted versions of the entire face. On the other hand the bases of NMF are localized features that correspond better to the intuitive notions of face parts [2]. An extension of NMF that gives even more localized bases by imposing additional locality constraints is the so-called LNMF [3].

NMF variants for object recognition have been proposed in [5, 6]. Various distance metrics suitable to NMF representation space have been proposed in [7]. Methods for initializing the weights and the bases of the NMF decomposition have been proposed in [8]. Theoretical aspects regarding why NMF gives a unique decomposition of an object into its parts are provided in [9].

In the proposed technique we incorporate discriminant constraints inside the NMF decomposition and that way a part based decomposition with enhanced discriminant power is taken. The introduced method results to a class specific decomposition that is unique for each facial (person) class. The intuitive motivation behind the class-specific methods is to find for every face a unique decomposition into its own discriminant parts. Class-specific discriminant transforms have been also used for discriminant dimensionality reduction in the feature vectors of the elastic grids and for discriminant weighting of their nodes [10–13]. The introduced algorithm is applied to the frontal face verification problem using the XM2VTS database.

## 2 Frontal Face Verification and Subspace Techniques

Let  $\mathcal{U}$  be a facial image database. Each facial image  $\mathbf{x} \in \mathcal{U}$  is supposed to belong to one of the  $K$  facial (person) classes  $\{\mathcal{U}_1, \mathcal{U}_2, \dots, \mathcal{U}_K\}$  with  $\mathcal{U} = \bigcup_{i=1}^K \mathcal{U}_i$ . For a face verification system that uses the database  $\mathcal{U}$ , a genuine (or client) claim is performed when a person  $t$  provides its facial image  $\mathbf{x}$ , claiming that  $\mathbf{x} \in \mathcal{U}_r$  and  $t = r$ . When a person  $t$  provides its facial image  $\mathbf{x}$  and claims that  $\mathbf{x} \in \mathcal{U}_r$ , with  $t \neq r$ , an impostor claim occurs. The scope of a face verification system is to handle properly these claims by accepting the genuine claims and rejecting the impostor ones.

Let the facial image database  $\mathcal{U}$  be comprised by  $L$  facial images  $\mathbf{x}_j \in \mathfrak{R}_+^F$ , where  $\mathfrak{R}_+ = [0, +\infty)$  and let the cardinality of each facial class  $\mathcal{U}_r$  to be  $N_r$ . A linear subspace transformation of the original  $F$ -dimensional space onto a  $M$ -dimensional subspace (usually  $M \ll F$ ) is a matrix  $\mathbf{W} \in \mathfrak{R}^{M \times F}$  estimated using the database  $\mathcal{U}$ . The new feature vector  $\hat{\mathbf{x}} \in \mathfrak{R}^M$  is given by:

$$\hat{\mathbf{x}} = \mathbf{W}\mathbf{x}. \quad (1)$$

The rows of the matrix  $\mathbf{W}$  contain the bases of the lower dimension feature space. The bases matrix  $\mathbf{W}$  could be the same for all facial classes of the database or could be unique for each facial class. In case of class-specific image bases, for the reference person  $r$ , the set  $\mathcal{I}_r = \mathcal{U} - \mathcal{U}_r$ , that corresponds to impostor images is used in order to construct the two-class problem (genuine versus impostor class) [11].

After the projection given by (1), a distance metric is chosen in order to measure the similarity of a test facial image to a certain class. This similarity measure can be the  $L_1$  norm, the  $L_2$  norm, the normalized correlation or the Mahalanobis distance. In case of face verification, the algorithm should also learn a threshold on the similarity measure in order to accept or reject a client/impostor claim.

### 3 The NMF algorithm

In order to apply NMF, the matrix  $\mathbf{X} \in \mathfrak{R}_+^{F \times L} = [x_{i,j}]$  should be constructed, where  $x_{i,j}$  is the  $i$ -th element of the  $j$ -th image. In other words the  $j$ -th column of  $\mathbf{X}$  is the  $\mathbf{u}_j$  facial image. NMF aims to find two matrices  $\mathbf{Z} \in \mathfrak{R}_+^{F \times M} = [z_{i,k}]$  and  $\mathbf{H} \in \mathfrak{R}_+^{M \times L} = [h_{k,j}]$  such that,

$$\mathbf{X} \approx \mathbf{ZH}. \quad (2)$$

The facial image  $\mathbf{x}_j$  after the NMF decomposition can be written as  $\mathbf{x}_j \approx \mathbf{ZH}_j$ , where  $\mathbf{h}_j$  is the  $j$ -th column of  $\mathbf{H}$ . Thus, the lines of the matrix  $\mathbf{Z}$  can be considered as bases images and the  $\mathbf{h}_j$  as the weight vector. The  $\mathbf{h}_j$  vectors can also be considered as the projected vectors of a lower dimensional feature space.

The NMF imposes non-negative constraints in both the elements of  $z_{i,k}$  and of  $h_{k,j}$ . Thus, only non-subtractive combinations are allowed. This is believed to correspond better to the intuitive notion of combining parts of face in order to create a whole one.

One of the algorithms initially proposed for finding the matrices  $\mathbf{Z}$  and  $\mathbf{H}$  used the Kullback-Leibler divergence [14]:

$$D_N(\mathbf{X}||\mathbf{ZH}) = \sum_{i,j} (x_{i,j} \ln(\frac{x_{i,j}}{\sum_k z_{i,k} h_{k,j}}) + \sum_k z_{i,k} h_{k,j} - x_{i,j}) \quad (3)$$

as the measure of the cost for factoring  $\mathbf{X}$  into  $\mathbf{ZH}$  [14]. The NMF factorization is the outcome of the optimization:

$$\begin{aligned} \min_{\mathbf{Z}, \mathbf{H}} D_N(\mathbf{X}||\mathbf{ZH}) \text{ subject to} \quad (4) \\ z_{i,k} \geq 0, h_{k,j} \geq 0, \sum_i z_{i,j} = 1, \forall j. \end{aligned}$$

By using an auxiliary function and the Expectation Maximization (EM) algorithm [14], the following update rules for  $h_{k,j}$  and  $z_{i,k}$  guarantee a non increasing behavior of (3). The update rule for the  $t$ -th iteration for  $h_{k,j}$  is given by:

$$h_{k,j}^{(t)} = h_{k,j}^{(t-1)} \frac{\sum_i z_{i,k}^{(t-1)} \frac{x_{i,j}}{\sum_l z_{i,l}^{(t-1)} h_{l,j}^{(t-1)}}}{\sum_i z_{i,k}^{(t-1)}} \quad (5)$$

whereas, for the  $z_{i,k}$ , the update rule is given by:

$$z_{i,k}^{(t)} = z_{i,k}^{(t-1)} \frac{\sum_j h_{k,j}^{(t)} \frac{x_{i,j}}{\sum_l z_{i,l}^{(t-1)} h_{l,j}^{(t)}}}{\sum_j h_{k,j}^{(t)}}. \quad (6)$$

Since  $\mathbf{x}_j \approx \mathbf{ZH}_j$ , a natural way to compute the projection of  $\mathbf{x}_j$  to a lower dimensional feature space using NMF is  $\hat{\mathbf{x}}_j = \mathbf{Z}^\dagger \mathbf{x}_j$ . The pseudo-inverse  $\mathbf{Z}^\dagger$  can be calculated using singular value decomposition methods [15]. In any case, we can not use the coefficient matrix  $\mathbf{H}$  computed directly from the update rules (which gives us its values in the training phase), since we do not have any expression for calculating this representation for the test images.

## 4 The LNMF algorithm

The idea of NMF decomposition was further extended to the LNMF [3] where additional constraints concerning the spatial locality of the bases were employed in the optimization problem defined in (4).

Let  $\mathbf{U} = [u_{i,j}] = \mathbf{Z}^T \mathbf{Z}$ ,  $\mathbf{V} = [v_{i,j}] = \mathbf{H} \mathbf{H}^T$ , both being  $M \times M$ , LNMF aims at learning local features by imposing the following three additional locality constraints on the NMF. The first constraint is to create bases that cannot be further decomposed into more components [3]. This is accomplished by making the bases as sparse as possible by imposing  $\sum_i u_{i,i}$  to be minimal [3].

Another constraint is to make the bases to be as orthogonal as possible, so as to minimize the redundancy between different bases. This can be imposed by requiring  $\sum_{i \neq j} u_{i,j}$  to be minimal. Another employed constraint, requires that  $\sum_i v_{i,i}$  is maximized [3].

When the above constraints are incorporated in (3), a new cost function is created as:

$$D_L(\mathbf{X}||\mathbf{ZH}) = D_N(\mathbf{X}||\mathbf{ZH}) + \alpha \sum_{i,j} u_{i,j} - \beta \sum_i v_{i,i} \quad (7)$$

where  $\alpha, \beta > 0$  are constants. A solution for the minimization of the cost given in (7) subject to non-negative constraints, can be found in [3]. In order to ensure that the cost function (7) is nonincreasing, the following update rules for  $z_{i,k}$  and  $h_{k,j}$  are employed:

$$h_{k,j}^{(t)} = \sqrt{h_{k,j}^{(t-1)} \sum_i z_{i,k}^{(t-1)} \frac{x_{i,j}}{\sum_l z_{i,l}^{(t-1)} h_{l,j}^{(t-1)}}} \quad (8)$$

$$\hat{z}_{i,k}^{(t)} = z_{i,k}^{(t-1)} \frac{\sum_j h_{k,j}^{(t)} \frac{x_{i,j}}{\sum_l z_{i,l}^{(t-1)} h_{l,j}^{(t)}}}{\sum_j h_{k,j}^{(t)}} \quad (9)$$

$$z_{i,k}^{(t)} = \frac{\hat{z}_{i,k}^{(t)}}{\sum_l \hat{z}_{l,k}^{(t)}}. \quad (10)$$

## 5 The CSDNMF algorithm

In this Section discriminant constraints are integrated inside the cost function (3). The minimization procedure of the new cost function yields a *Class-Specific Discriminant Non-negative Matrix Factorization* (CSDNMF) method. In order to formulate the CSDNMF decomposition, the facial image vectors of the genuine claims to the reference person  $r$  are in the first  $N_r = N_G$  columns of the matrix  $\mathbf{X}$ . Then, the columns from  $N_r + 1$  to  $L$  correspond to impostor claims. The total number of impostor claims is  $N_I = L - N_r$ . The coefficient vector  $\mathbf{h}_j$  of the image  $\mathbf{x}_j$  that corresponds to the  $\rho$ -th image of the genuine class will be denoted as  $\boldsymbol{\eta}_\rho^{(G)}$ . If the facial vector  $\mathbf{x}_j$  is the  $\rho$ -th image of the impostor class then the corresponding coefficient vector  $\mathbf{h}_j$  will be denoted as  $\boldsymbol{\eta}_\rho^{(I)}$ .

Let a distance metric (e.g. the  $L_2$  norm) be used in order to quantify the similarity of a test facial image vector  $\mathbf{x}_j$  to a given facial class. It sounds reasonable to require that the feature vectors corresponding to the genuine class, should have great similarity (small distance metric value), while the feature vectors of the impostor class should have small similarity (large distance metric value).

In order to define the similarity of the projection  $\mathbf{h}_j$  of the facial image  $\mathbf{x}_j$  to a given class  $r$  in the feature space of the coefficients, the  $L_2$  norm can be used as:

$$d_r(\mathbf{h}_j) = \|\mathbf{h}_j - \boldsymbol{\mu}^{(G)}\|^2 \quad (11)$$

where  $\boldsymbol{\mu}^{(G)}$  is the mean vector of the vectors  $\boldsymbol{\eta}_\rho^{(G)}$ . In the reduced feature space of the vectors  $\mathbf{h}_j$  we demand that the similarity measures  $d_r(\boldsymbol{\eta}_\rho^{(I)})$  (impostor similarity measures) to be maximized while minimizing the similarity measures  $d_r(\boldsymbol{\eta}_\rho^{(G)})$  (genuine similarity measures). Then the optimization problem for the class  $r$  is the maximization of:

$$\frac{1}{N_I} \sum_{\mathbf{x}_j \in \mathcal{I}_r} d_r(\mathbf{h}_j) = \frac{1}{N_I} \sum_{\rho=1}^{N_I} \|\boldsymbol{\eta}_\rho^{(I)} - \boldsymbol{\mu}^{(G)}\|^2 = \text{tr}[\mathbf{W}_r], \quad (12)$$

where  $\mathbf{W}_r = \frac{1}{N_I} \sum_{\rho=1}^{N_I} (\boldsymbol{\eta}_\rho^{(I)} - \boldsymbol{\mu}^{(G)})(\boldsymbol{\eta}_\rho^{(I)} - \boldsymbol{\mu}^{(G)})^T$ . The second optimization problem is the minimization of:

$$\frac{1}{N_G} \sum_{\mathbf{x}_j \in \mathcal{U}_r} d_r(\mathbf{h}_j) = \frac{1}{N_G} \sum_{\rho=1}^{N_G} \|\boldsymbol{\eta}_\rho^{(G)} - \boldsymbol{\mu}^{(G)}\|^2 = \text{tr}[\mathbf{B}_r], \quad (13)$$

where  $\mathbf{B}_r = \frac{1}{N_G} \sum_{\rho=1}^{N_G} (\boldsymbol{\eta}_\rho^{(G)} - \boldsymbol{\mu}^{(G)})(\boldsymbol{\eta}_\rho^{(G)} - \boldsymbol{\mu}^{(G)})^T$ . We impose these two additional constraints in the cost function given in (4) as:

$$D_c(\mathbf{X} \|\mathbf{Z}_r \mathbf{H}_r) = D_N(\mathbf{X} \|\mathbf{Z}_r \mathbf{H}_r) + \zeta \text{tr}[\mathbf{B}_r] - \theta \text{tr}[\mathbf{W}_r] \quad (14)$$

where  $\zeta, \theta > 0$  are constants. The minimization of (14) gives a person specific decomposition (different bases  $\mathbf{Z}_r$  for each reference face class  $r$ ).

In order to derive the coefficients of CSDNMF we have used an auxiliary function similar to those used in the EM algorithm in [14]. Let  $G$  be an auxiliary function for  $Y(\mathbf{F})$  if  $G(\mathbf{F}, \mathbf{F}^{(t-1)}) \geq Y(\mathbf{F})$  and  $G(\mathbf{F}, \mathbf{F}) = \mathbf{F}$ . If  $G$  is an auxiliary function of  $Y$ , then  $Y$  is nonincreasing under the update  $\mathbf{F}^t = \arg \min_{\mathbf{F}} G(\mathbf{F}, \mathbf{F}^{(t-1)})$ . Let  $r$  be the reference facial class, we can prove that  $\mathbf{G}_c(\mathbf{H}, \mathbf{H}^{(t-1)})$  is an auxiliary function of  $Y_c(\mathbf{H}) = D_c(\mathbf{X} \|\mathbf{Z}_r \mathbf{H}_r)$ , where  $\mathbf{G}_c(\mathbf{H}, \mathbf{H}^{(t-1)})$  is given by:

$$\begin{aligned} \mathbf{G}_c(\mathbf{H}, \mathbf{H}^{(t-1)}) &= \sum_i \sum_j (x_{i,j} \ln x_{i,j} - x_{i,j}) + \\ &+ \sum_i \sum_j \sum_k \frac{z_{i,k} h_{k,j}^{(t-1)}}{\sum_l z_{i,l} h_{l,j}^{(t-1)}} (\ln(z_{i,k} h_{k,j}) - \ln \frac{z_{i,k} h_{k,j}^{(t-1)}}{\sum_l z_{i,l} h_{l,j}^{(t-1)}}) + \\ &+ \sum_i \sum_j \sum_k z_{i,k} h_{k,j} + \zeta \text{tr}[\mathbf{B}_r] - \theta \text{tr}[\mathbf{W}_r]. \end{aligned} \quad (15)$$

It is straightforward to show that  $\mathbf{G}_c(\mathbf{H}, \mathbf{H}) = Y_c(\mathbf{H})$ . In order to prove that  $\mathbf{G}_c(\mathbf{H}, \mathbf{H}^{(t-1)}) \geq Y_c(\mathbf{H})$  since,  $\ln(\sum_k z_{i,k} h_{k,j})$  is convex, the following inequality

holds:

$$-\ln\left(\sum_k z_{i,k} h_{k,j}\right) \leq -\sum_k a_k \ln \frac{z_{i,k} h_{k,j}}{a_k} \quad (16)$$

for all non-negative  $a_k$  that satisfy  $\sum_k a_k = 1$ . By letting  $a_k = \frac{z_{i,k} h_{k,j}^{(t-1)}}{\sum_l z_{i,l} h_{l,j}^{(t-1)}}$  we obtain:

$$-\ln\left(\sum_k z_{i,k} h_{k,j}\right) \leq \sum_k \frac{z_{i,k} h_{k,j}^{(t-1)}}{\sum_l z_{i,l} h_{l,j}^{(t-1)}} \left( \ln(z_{i,k} h_{k,j}) - \ln \frac{z_{i,k} h_{k,j}^{(t-1)}}{\sum_l z_{i,l} h_{l,j}^{(t-1)}} \right). \quad (17)$$

From (17) it is straightforward to show that  $\mathbf{G}_c(\mathbf{H}, \mathbf{H}^{(t-1)}) \geq Y_c(\mathbf{H})$ . Thus,  $\mathbf{G}_c(\mathbf{H}, \mathbf{H}^{(t-1)})$  is an auxiliary function of  $Y_c(\mathbf{H})$ .

In this decomposition we have two different update rules. One for the genuine class and one for the impostor class. For  $l = 1, \dots, N^G$  (genuine class) the update rules for the coefficients  $h_{k,l}$  for the reference person  $r$  are given by letting  $\frac{\partial G_c(\mathbf{H}, \mathbf{H}^{(t-1)})}{\partial h_{k,l}} = 0$ . Then,

$$\begin{aligned} \frac{\partial G_c(\mathbf{H}, \mathbf{H}^{(t-1)})}{\partial h_{k,l}} &= -\sum_i x_{i,l} \frac{z_{i,k} h_{k,l}^{(t-1)}}{\sum_n z_{i,n} h_{n,l}^{(t-1)}} \frac{1}{h_{k,l}} + \sum_i z_{i,k} + \\ &+ 2\zeta(h_{k,l} - \mu_k^{(G)}) \frac{1}{N^G} - 2\theta(\mu_k^{(G)} - \mu_k^{(I)}) \frac{1}{N^G} = 0. \end{aligned} \quad (18)$$

The quadratic equation (18) is expanded as:

$$\begin{aligned} -\sum_i x_{i,l} \frac{z_{i,k} h_{k,l}^{(t-1)}}{\sum_n z_{i,n} h_{n,l}^{(t-1)}} + (1 - (2\zeta + 2\theta)) \frac{1}{N^G} \left( \frac{1}{N^G} \sum_{\lambda, \lambda \neq l} h_{k,\lambda} \right) + 2\theta \frac{1}{N^G} \mu_k^{(I)} h_{k,l} + \\ + \frac{1}{N^G} (2\zeta - (2\zeta + 2\theta)) \frac{1}{N^G} h_{k,l}^2 = 0. \end{aligned} \quad (19)$$

By solving the quadratic equation (19) the update rules for the  $h_{k,l}$  of the genuine class are:

$$h_{k,l} = \frac{T + \sqrt{T^2 + 4 \frac{1}{N^G} (2\zeta - (2\zeta + 2\theta)) \frac{1}{N^G} h_{k,l}^{(t-1)} \sum_i z_{i,k}^{(t-1)} \frac{x_{i,j}}{\sum_n z_{i,n}^{(t-1)} h_{n,l}^{(t-1)}}}}{2 \frac{1}{N^G} (2\zeta - (2\zeta + 2\theta)) \frac{1}{N^G}} \quad (20)$$

where  $T$  is given by:

$$T = (2\zeta + 2\theta) \frac{1}{N^G} \left( \frac{1}{N^G} \sum_{\lambda, \lambda \neq l} h_{k,\lambda} \right) - 2\theta \frac{1}{N^G} \mu_k^{(I)} - 1. \quad (21)$$

The update rules for the coefficients  $h_{k,l}$  for the impostor class of the reference person  $r$  are given by letting  $\frac{\partial G_c(\mathbf{H}, \mathbf{H}^{(t-1)})}{\partial h_{k,l}} = 0$ :

$$\frac{\partial G_c(\mathbf{H}, \mathbf{H}^{(t-1)})}{\partial h_{k,l}} = -\sum_i x_{i,l} \frac{z_{i,k} h_{k,l}^{(t-1)}}{\sum_n z_{i,n} h_{n,l}^{(t-1)}} \frac{1}{h_{k,l}} + \sum_i z_{i,k} - 2 \frac{1}{N^I} \theta (h_{k,l} - \mu_k^{(G)}) = 0 \quad (22)$$

where  $j = N_G + 1, \dots, L$ . By solving the quadratic equation (22) the update rules for the  $h_{k,l}$  are given by:

$$h_{k,l} = \frac{2\theta\mu_k^{(G)} + N_I + \sqrt{(2\theta\mu_k^{(G)} + N_I)^2 - 8N_I\theta h_{k,l}^{(t-1)} \sum_i z_{i,k}^{(t-1)} \frac{x_{i,j}}{\sum_n z_{i,n}^{(t-1)} h_{n,l}^{(t-1)}}}}{4\theta}. \quad (23)$$

It can be easily proven that the update rules for the bases matrix  $\mathbf{Z}_r = [z_{i,k}]$  for the reference person  $r$  are given by:

$$z_{i,k}^{(t)} = z_{i,k}^{(t-1)} \frac{\sum_j h_{k,j}^{(t)} \frac{x_{i,j}}{\sum_l z_{i,l}^{(t-1)} h_{l,j}^{(t)}}}{\sum_j h_{k,j}^{(t)}} \quad (24)$$

and

$$z_{i,k}^{(t)} = \frac{z_{i,k}^{(t)}}{\sum_l z_{l,k}^{(t)}}. \quad (25)$$

When someone claims that a test image  $\mathbf{x}$  corresponds to a reference facial class  $r$ , then  $\mathbf{x}$  is projected using the  $\mathbf{Z}_r^\dagger$  matrix as  $\hat{\mathbf{x}} = \mathbf{Z}_r^\dagger \mathbf{x}$ .

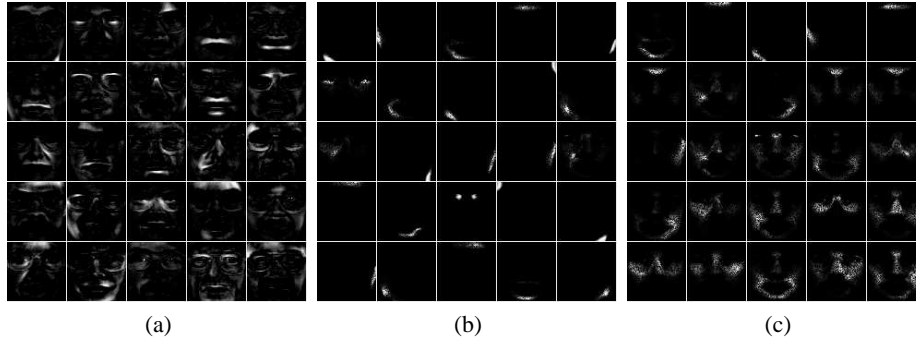
## 6 Experimental Results

The experiments were conducted in the XM2VTS database using the protocol described in [16]. The images were aligned semi-automatically according to the eyes position of each facial image using the eye coordinates. The facial images were down-scaled to  $64 \times 64$  resolution. Histogram equalization was used for normalizing the facial images. The XM2VTS database provides two experiment setups namely, Configuration I and Configuration II [16]. Each Configuration is divided in three different sets the training set, the evaluation set and the test set. The training set is used to create client and impostor models for each person. The evaluation set is used to learn the thresholds.

The training set of the Configuration I contains 200 persons with 3 images per person. The evaluation set contains 3 images per client for genuine claims and 25 evaluation impostors with 8 images per impostor. Thus, evaluation set gives a total of  $3 \times 200 = 600$  client claims and  $25 \times 8 \times 200 = 40.000$  impostor claims. The test set has 2 images per client and 70 impostors with 8 images per impostor and gives  $2 \times 200 = 400$  client claims and  $70 \times 8 \times 200 = 112.000$  impostor claims. In the training set the matrices of the basis images for NMF and LNMF decompositions are learned. These matrices are common for all persons. In case of CSNMF the training set is used for calculating for each reference person  $r$  a different set of bases for feature selection. For visual comparison a number of 25 images for the NMF, the LNMF and the proposed CSDNMF (for the first person in the training set) are given in Figure 1.

The facial images have been projected using these bases into a low dimensional feature space and the normalized correlation was used in order to define the similarity measure between two faces as:

$$D(\mathbf{x}_r, \mathbf{x}_t) = \frac{\hat{\mathbf{x}}_r^T \hat{\mathbf{x}}_t}{\|\hat{\mathbf{x}}_r\| \|\hat{\mathbf{x}}_t\|} \quad (26)$$



**Fig. 1.** A set of 25 bases images for (a) NMF, (b) LNMF and (c) CSDNMF.

where  $\mathbf{x}_r$  and  $\mathbf{x}_t$  are the reference and the test facial image respectively, while  $\hat{\mathbf{x}}_r$  and  $\hat{\mathbf{x}}_t$  are their projections to one of the subspace.

The similarity measures for each person, calculated in both evaluation and training set form the distance vector  $\mathbf{d}(r)$ . The elements of the vector  $\mathbf{d}(r)$  are sorted in descending order and are used for the person specific thresholds on the distance measure. Let  $T_Q(r)$  denote the  $Q$ -th order statistic of the vector of distances,  $\mathbf{d}(r)$ . The threshold of the person  $r$  is chosen to be equal to  $T_Q(r)$ . Let  $\mathbf{x}_r^1$ ,  $\mathbf{x}_r^2$  and  $\mathbf{x}_r^3$  be the 3 instances of the person  $r$  in the training set. A claim of a person (with a facial image  $\mathbf{x}_t$ ) to the identity  $r$  is considered valid if  $\max_j \{D(\mathbf{x}_r^j, \mathbf{x}_t)\} < T_Q(r)$ . Obviously when varying  $Q$ , different pairs of *False Acceptance Rate* and *False Rejection Rate* can be created and that way a *Receiver Operating Characteristic* (ROC) curve is produced and the *Equal Error Rate* (EER) can be measured [11, 16].

The performance of the NMF, LNMF and CSDNMF algorithms for various feature dimensions in the test set of Configuration I is illustrated in Figure 2. The best EER achieved for CSDNMF is 3.4% when more than 110 dimensions are kept. The best EER for NMF and LNMF is more than 8%. That is, a decrease of more than 4% in terms of EER has been achieved.

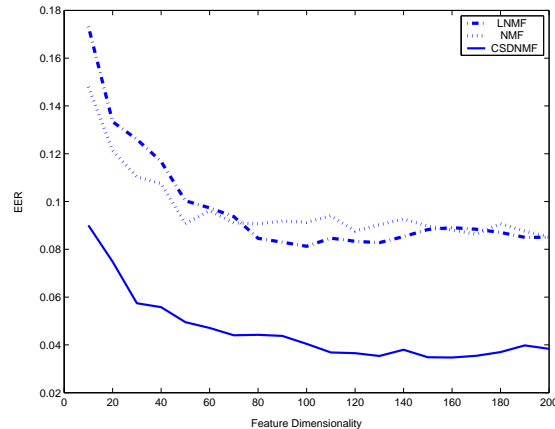
## 7 Conclusions

We incorporated discriminant constraints in the cost of NMF decomposition in order to extract class-specific discriminant non-negative decomposition. We solved the new optimization problem by developing update rules for both the weighting coefficients and the bases. We applied the new decomposition to frontal face verification where better performance than NMF and LNMF has been achieved.

## Acknowledgment

This work is funded by the integrated project BioSec IST-2002-001766 (Biometric Security, <http://www.biosec.org>), under Information Society Technologies (IST) priority of the 6th Framework Programme of the European Community.





**Fig. 2.** EER as a function of feature dimension.

## References

1. M. Kirby and L. Sirovich, "Application of the karhunen-loeve procedure for the characterization of human faces.," *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 103–108, Jan. 1990.
2. D.D. Lee and H.S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, pp. 788–791, 1999.
3. S.Z. Li, X.W. Hou, and H.J. Zhang, "Learning spatially localized, parts-based representation," in *CVPR*, Kauai, HI, USA, December 8-14 2001, pp. 207–212.
4. P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection.," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, July 1997.
5. D. Guillamet, J. Vitria, and B. Schiele, "Introducing a weighted non-negative matrix factorization for image classification," *Pattern Recognition Letters*, vol. 24, no. 14, pp. 2447 – 2454, 2003.
6. L. Weixiang and N. Zheng, "Non-negative matrix factorization based methods for object recognition," *Pattern Recognition Letters*, vol. 25, no. 9-10, pp. 893–897, 2004.
7. D. Guillamet and Vitria, "Evaluation of distance metrics for recognition based on non-negative matrix factorization," *Pattern Recognition Letters*, vol. 24, no. 9-10, pp. 1599 – 1605, 2003.
8. S. Wild, J. Curry, and A. Dougherty, "Improving non-negative matrix factorizations through structured initialization," *Pattern Recognition*, vol. 37, pp. 2217–2232, 2004.
9. D. Donoho and V. Stodden, "When does non-negative matrix factorization give a correct decomposition into parts ?," *Advances in Neural Information Processing Systems*, vol. 17, 2004.
10. A. Tefas, C. Kotropoulos, and I. Pitas, "Face verification using elastic graph matching based on morphological signal decomposition," *Signal Processing*, vol. 82, no. 6, pp. 833–851, 2002.
11. A. Tefas, C. Kotropoulos, and I. Pitas, "Using support vector machines to enhance the performance of elastic graph matching for frontal face authentication," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 7, pp. 735–746, 2001.

12. C. Kotropoulos, A. Tefas, and I. Pitas, "Frontal face authentication using discriminating grids with morphological feature vectors.," *IEEE Transactions on Multimedia*, vol. 2, no. 1, pp. 14–26, Mar. 2000.
13. C. Kotropoulos, A. Tefas, and I. Pitas, "Frontal face authentication using morphological elastic graph matching.," *IEEE Transactions on Image Processing*, vol. 9, no. 4, pp. 555–560, Apr. 2000.
14. D.D. Lee and H.S. Seung, "Algorithms for non-negative matrix factorization," in *NIPS*, 2000, pp. 556–562.
15. G.H. Golub and C.F. VanLoan, *Matrix Computations*, third ed. John Hopkins Univ. Press, 1996.
16. K. Messer, J. Matas, J.V. Kittler, J. Luetten, and G. Maitre, "Xm2vtsdb: The extended m2vts database," in *AVBPA'99*, 1999, pp. 72–77.