Face Recognition with Non-negative Matrix Factorization

Menaka Rajapakse and Lonce Wyse

Institute for Infocomm Research
21 Heng Mui Keng Terrace, Singapore 119613

ABSTRACT

A face can conceptually be represented as a collection of sparsely distributed parts: eyes, nose, mouth, etc. We use Non-negative Matrix Factorization (NMF) to yield sparse representation of localized features to represent distributed parts over a human face. This paper explores the potential of NMF for face recognition and the possibilities for gender-based features in face reconstruction. Further, we compare the results of NMF with other common face recognition methods.

Keywords: NMF, Gabor, PCA, Face Recognition

1. INTRODUCTION

Due to the redundancies and overlapping of features in high dimensions, analyzing objects in data space can be computationally exhaustive. Therefore, object recognition is often converted to a problem in a low dimensional subspace. A subspace representation also exposes hidden structures and characteristics inherent to the type of objects. Furthermore, this subspace must be able to effectively cope with object features that are amenable to lighting and other variations. Hence, the selection of an appropriate reduced data representation is imperative for the success and efficacy of an object recognition system.

Linear basis methods such as PCA[1] and factor analysis are often used as representations for multivariate data. With PCA, a global representation of the data is achieved and the data reduction is accomplished by keeping only the components which account for the maximum variance of data and eliminating components with smaller eigen values. By choosing the directions of projections, which maximize total scatter, PCA tend to retain unwanted variations due to lighting and facial expression.[2] Moreover, though PCA linear projections are optimal in a least square sense for reconstruction from a low dimensional basis, their abilities to discriminate a given set of objects are non-optimal. Another data-adaptive approach known as sparse coding has given fresh insight into data representation.[3] In sparse coding, an image is represented by a small number of *active* feature components. The distribution of weights for the feature vectors peaks at 0 and has only a few features with significant weights for most reconstructions.[3] Olshausen and Field[4] showed that there exist similarities between the features derived from linear sparse coding of natural images and the receptive fields of *simple cells* in the visual cortex. Natural images are often assumed to have a sparse structure based on the fact that they can be described with localized features such as edges, lines and other descriptive elements[3]; this is especially applicable to face images.

A multivariate data reduction approach known as Non-negative Matrix Factorization (NMF) has recently attracted significant attention in part because it is suggestive of some aspects of activation patterns in response to images in the mammalian visual cortex. In NMF, as the name implies, the non-negativity adds constraints to the matrix factorization, allowing only additions in the synthesis; there are no cancellations or interference of patterns via subtraction or negative feature vector values. This more naturally leads to the notion of parts-based representation of images[5].[6] With the underlying non-negative constraints, NMF is able to learn localized parts based representations. Sparse coding with NMF seems befitting especially for face recognition applications as the features of face images are naturally represented as a small collection of features, namely eyes, nose and mouth, which are distributed over the face. Because the outputs of NMF are localized features, we can use these parts

---

Further author information: (Send correspondence to Menaka)
Menaka: E-mail: menaka@i2r.a-star.edu.sg
Lonce: E-mail: lonce@i2r.a-star.edu.sg

based features collectively to represent a face. Similar to the eigenface approach where a face is represented by linearly combined basis faces, in NMF a face is represented by a linear combination of basis images of localized features. For face recognition applications, the drawbacks of using eigenfaces include sensitivity to intensity changes, and translations. Furthermore, classes are often not very compactly grouped in this feature space.

In contrast, Gabor wavelet based techniques are less sensitive to lighting and scale variations[7][8].[9] A Gabor wavelet technique known as elastic bunch graph matching[8] has been effective in face recognition applications, but the recognition procedure involves complex computations. Another approach,[9] uses a set of feature points representing relative locations of facial landmarks points to extract Gabor features at different scales and orientations. The final performance of this system relies on the degree of accuracy of feature point localization.

In contrast to the eigenface approach where no perceptually significant features are detected, both the Gabor based and NMF based methods make use of such features. Gabor wavelets exhibit properties that are similar to the those of *edge detecting* receptive fields of simple cells in the visual cortex.[10] Some recent findings support the idea of parts based representation in the brain, according to which, the brain utilizes a lexicon of parts to represent and recognize shapes. *"The number of different objects that we have to represent in the brain and store in memory is virtually infinite. And you have to do that with a large but finite number of neurons, so a parts-based approach makes sense. Represent something in terms of its parts, and you've got a really high number of combinations, just as letters of the alphabet encode thousands of words".*[11] Thus the part-based features extracted with NMF have biological motivation suggestive of their suitability for object recognition.

Recently, Xian et al.,[12] have proposed a variant of NMF known as local non-negative matrix factorization (LNMF). In LNMF, a localization constraint is imposed in addition to the NMF's inherent sparseness so that the learned parts are an over complete set of even more spatially localized features. According to their results, LNMF yields better recognition accuracy for the ORL database compared to NMF.

In this paper, we further analyze the use of NMF for face recognition by a comparative study with well-known eigenface and Gabor wavelet based techniques, and then extends it to assess more perceptual categories of faces, such as gender representations. This paper is organized as follows. In section 2, we will review the basic concepts of NMF. Face recognition with NMF encoding is addressed in Section 3. Section 4 will present some experiments and results and we will compare some of the existing face recognition methods discussed above with NMF based approach. Finally, section 5 will summarize the conclusions of the paper.

2. NON-NEGATIVE MATRIX FACTORIZATION

Given a data matrix $F = \{F_{ij}\}_{nxm}$, non-negative matrix factorization refers to the decomposition of the matrix F into two matrices W and H of size nxr and rxm, respectively, such that

$$F = WH \tag{1}$$

where the elements in W and H are all positive values. From this decomposition, a reduced representation is achieved by choosing r such that $r < n$ and $r < m$.

In NMF, no negative values are allowed in matrix factors W and H. The non-negativity constraint is imposed in factorizing the data matrix F by limiting data manipulation only to additions; no subtractions are allowed. The reconstruction of an object is performed only by adding its representative parts collectively. Each column in the matrix W is called a basis image, and a column in the matrix H is called an encoding. An image (column) in F can be reconstructed by linearly combining basis images with the coefficients in an encoding. The encodings influence the activation of pixels in the original matrix via basis images.

Given a data matrix F, Lee and Seung[5] found a technique for factorizing the F to yield matrices W and H as given in Eq(1). Each element in the matrix F can be written as $F_{ij} = \sum_{\rho=1}^{r} W_{i\rho} H_{\rho j}$ where r represents the number of basis images and the number of coefficients in an encoding. The following iterative learning rules are used to find the linear decomposition[5]:

$$H_{\rho j} \leftarrow H_{\rho j} \sum_{i=1}^{n} (\frac{W_{i\rho} F_{ij}}{\sum_{k=1}^{r} W_{ik} H_{kj}}) \tag{2}$$

$$W_{i\rho} \leftarrow W_{i\rho} \sum_{j=1}^{m} (\frac{F_{ij}H_{\rho j}}{\sum_{k=1}^{r} W_{ik}H_{kj}}) \qquad (3)$$

$$W_{i\rho} \leftarrow \frac{W_{i\rho}}{\sum_{k=1}^{n} W_{k\rho}} \qquad (4)$$

The above *unsupervised* multiplicative learning rules are used iteratively to update W and H. The initial values of W and H are fixed randomly. At each iteration, a new value for W or H is evaluated. Each update consists of a multiplication and sums of positive factors. With these iterative updates, the quality of the approximation of the Eq.(1) improves monotonically with a guaranteed convergence to a locally optimal matrix factorization.[6]

3. NMF FOR FACE RECOGNITION

3.1. Representation and Training

The data matrix F, is constructed such that the training face images occupy the columns of the F matrix. Let the set of faces be $\Gamma = \{\mathbf{f}_1, \mathbf{f}_2, \cdots, \mathbf{f}_m\}$, then the data matrix, $F = [\mathbf{f}_1 \mathbf{f}_2 \cdots \mathbf{f}_m]$. Now learning is done using Eqs (2)-(4) to decompose the matrix F into two matrices, H and W. Let the basis images be $W = [\mathbf{w}_1 \mathbf{w}_2 \cdots \mathbf{w}_r]$ and encodings be $H = [\mathbf{h}_1 \mathbf{h}_2 \cdots \mathbf{h}_m]$. Each face $\mathbf{f}_i$ in F can be approximately reconstructed by linearly combining the basis images, and the corresponding encoding coefficients $\mathbf{h}_i = (h_{1i}h_{2i} \cdots h_{ri})^{\mathrm{T}}$ as shown in Figure 1. Hence, a face can be modeled in terms of a linear superposition of basis functions together with encodings as follows:

$$\mathbf{f}_i = W\mathbf{h}_i \qquad (5)$$

For each face $\mathbf{f}_i$ in the training set and test set, we calculate the corresponding encoding coefficients. The basis images in W are generated from the set of training faces, Γ^{train}. The encodings, $\mathbf{h}_i$ of each training face $\mathbf{f}_i$ is given by

$$\mathbf{h}_i = W^{-1}\mathbf{f}_i$$

Once trained, the face image set, $\{\mathbf{f}_1, \mathbf{f}_2, \cdots, \mathbf{f}_m\}$ is represented by a set of encodings $\{\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_m\}$ with reduced dimension, r.

3.2. Testing

Given a mean corrected face image $\mathbf{f}$, we can find a representative encoding for $\mathbf{f}$ as follows:

$$\mathbf{h} = W^{-1}\mathbf{f}$$

Figure (2) illustrates an encoding of a face when the rank is equal to 64. A distance metric is used to calculate the similarity between encodings of a trained image $\mathbf{h}_i \in \Gamma^{train}$ and a test image $\mathbf{h} \in \Gamma^{test}$. The cosine of the angle between the two data vectors is taken as the similarity measure:

$$\mathbf{s}_i = \frac{\mathbf{h} \cdot \mathbf{h}_i}{|\mathbf{h}||\mathbf{h}_i|} \qquad (6)$$

The similarity measure $\mathbf{s}_i$ determines the matching score between the encodings $\mathbf{h}$ and $\mathbf{h}_i$ corresponding to 2 faces $\mathbf{f}$ and $\mathbf{f}_i$. The optimum matching encoding of a trained image can be given as $\mathbf{h}_{i*}$ where

$$i^* = \arg \max_i \mathbf{s}_i$$

and $\mathbf{s}_i > h_{thresh}$ indicating that the face $\mathbf{f}_i{}^*$ is identified as the closest match for the face $\mathbf{f}$. Hence, the best matched trained image for a given test image is the one that maximizes $\mathbf{s}_i$ provided that the score is above a threshold, h_{thresh}. If there are no $\mathbf{h}_{i*}$ for which the score is greater than the threshold, the image is *rejected*. The h_{thresh} is determined empirically as the point at which the false acceptance (misclassification) rate and false rejection rate are equal - *equal error rate*.

Figure 1. The far left image is the original face image, and the reconstructed image using basis images (W) and the corresponding image encoding ($\mathbf{h}_i$) is shown right to the original image.

Figure 2. Coefficients of a single face encoding (h) as a 8x8 image (r=64)

4. EXPERIMENTS

We investigate the performance of NMF for the Surrey face database and for distinguishing male and female images. As the database contains more male participants than female, we take all female participants and a subset of male images so as to equalize the influence from each gender category. The remaining set of male images is used to test the behavior of the technique on previously unseen images - *imposters*. The following sections describe the approach in detail.

4.1. Database

The experiments were carried out on M2VTS face database from University of Surrey, which consists of 1180 images, with 4 images per person taken at four different times (one month apart). Though similar lighting conditions and backgrounds have been used during image acquisition, significant changes in hair styles, facial hair, presence and absence of glasses introduce variability into the images. These images are of frontal and near frontal views with somewhat dissimilar facial expressions. The original image size is 726 x 576 pixels and the database contains images of Caucasians and Asian males and females. The images are normalized for scale, rotation, translation and illumination invariance.

4.2. Preprocessing and Normalization

The image preprocessing facilitates minimizing the variances among faces of the same individual while maximizing the variance between different individuals. These variances occur in scale, rotation, translation and lighting conditions existing among captured images at different time instances.

The normalized output images have similar grey levels across all images in the database. The eye positions are fixed at preset coordinates. For the experiments, the final image size is reduced to 64x64 from the normalized image size of 150x200 by re-sampling the images.

The geometric normalization used in our approach is based on the manually located eye positions. In order to achieve faces invariant to rotation, translation and scale, a transformation matrix is computed by joining the located eye positions on a horizontal segment having a length of 52 pixels separating the two eyes in the original dimensions. The re-distribution of intensity values of the image is carried out using histogram equalization thereby producing an image with equally distributed intensity values.

4.3. Data Preparation

Given a database B of m face images, we divide B into 3 subsets Γ_1, Γ_2, and γ where Γ_1 represents female images, Γ_2 male images and γ imposters. Therefore $B = \Gamma_1 \cup \Gamma_2 \cup \gamma$ in which $\Gamma_i \cap \Gamma_j = \emptyset$, $\Gamma_i \cap \gamma = \emptyset$ and $\Gamma_j \cap \gamma = \emptyset$. Each set Γ_1 and Γ_2 consists of 556 images for 139 individuals where each individual is represented by 4 images. The imposter set, γ consists of 68 images represented by 17 male individuals with 4 images per each individual.

Data preparation for single fold cross-validation is done by dividing $\Gamma_l = \Gamma_1 \cup \Gamma_2$ into 2 mutually exclusive and exhaustive sets: Γ_l^{train}, Γ_l^{test}. We can then express the validation sets as: $\{\Gamma_l^k; l = 1, 2; k = train, test\}$. Each individual contributes 3 face images for the training set and a single image to the test set except in the case of imposters. For imposters, all 4 images are used for testing as no training is involved with the imposter images.

4.4. Experiment 1: Database Reconstruction

Using the iterative rules given in Eqs. (2)-(4), we estimate W and H so that we can re-write the Eq(1) with the estimated weight matrix $\hat{W}$ and the estimated encodings $\hat{H}$ so that

$$\hat{F} = \hat{W}\hat{H}.$$

For a given face $\mathbf{f}$ in F, the estimated face in $\hat{F}$ can be written as $\hat{\mathbf{f}}$. The information loss from reconstruction can be expressed as:

$$\varepsilon(\mathbf{f}) = \| \mathbf{f} - \hat{\mathbf{f}} \|^2 \tag{7}$$

4.4.1. Evaluation of RMS Error, RMS SNR, and EER

One possible source of image reconstruction and classification error is that the iterative training may stop before convergence is achieved or the system may converge to a non-global optimum.[6] Another source of error comes from *data reduction* aspect of the basic factorization Eq.(1), where each of the m training images are being represented as a linear combination of r basis images, or *feature* vectors, with r chosen to be significantly less than the number of training images.

For each face image in the database, we calculate the root-mean-square error per pixel between f and $\hat{f}$. The average rms error e_{rms} for an image is calculated by averaging over the number of faces m in the database.

$$e_{\text{rms}} = \frac{\sum_{i=0}^{m-1} \left[\frac{1}{n} \sum_{x=0}^{n-1} \left[\hat{f}_i(x) - f_i(x) \right]^2 \right]^{1/2}}{m}$$

where n is the number of pixels in an image. We then estimate the average rms signal-to-noise ratio for a reconstructed face as shown below.

$$SNR_{\text{rms}} = \frac{\sum_{i=0}^{i=m-1} \left[\frac{\sum_{x=0}^{n-1} \hat{f}(x)^2}{\sum_{x=0}^{n-1} \left[\hat{f}_i(x) - f_i(x) \right]^2} \right]^{1/2}}{m}$$

Figures (3)-(a)(b) illustrate the average rms signal-to-noise ratio and the average grey level rms errors for increasing r for the Surrey database of 1180 images. According to Figure (3), we can see that with increasing r the average rms signal to noise ratio improves while the average rms error of pixel grey level decreases.

The choice of r represents a tradeoff between accuracy and data reduction. For all of our experiments, we chose an r equal to 64.

4.5. Experiment 2: Face Recognition with Cross-validation

The objective of this experiment is to measure the over all face recognition performance for the registered images in the database. This experiment is performed for NMF, PCA and Gabor methods and the results are compared for their recognition accuracy. For NMF, we also measure the rejection rate when imposters are presented to the system.

4.5.1. PCA Approach

For the experiment with PCA, we use 100 eigenfaces.[1] These eigenfaces are generated using 885 faces in the database and the first 100 eigenfaces corresponding to the largest eigenvalues were retained.
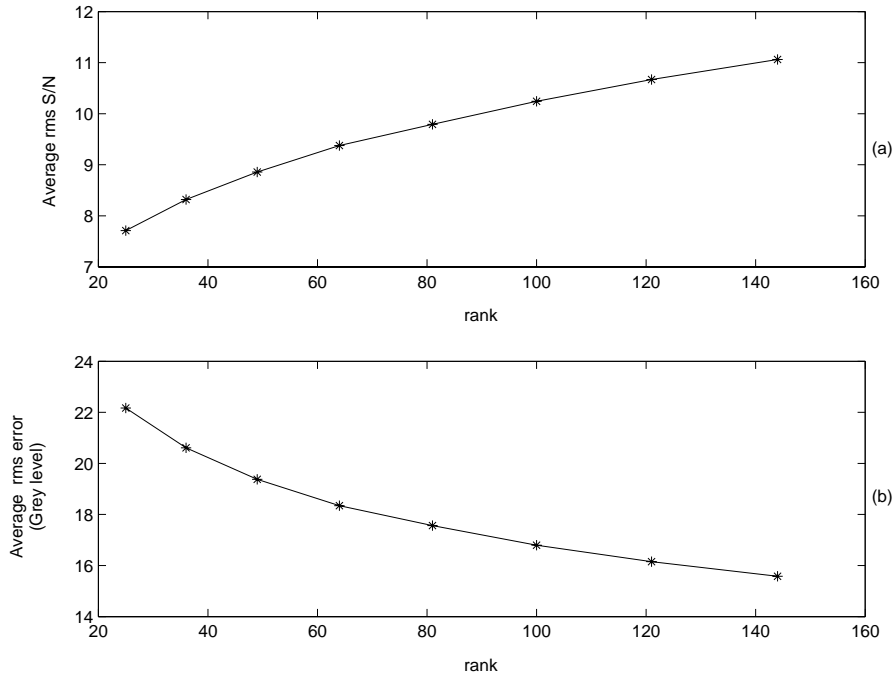
Figure 3. (a)Average rms signal-to-noise ratio (b)Average rms-error in grey level for increasing r (r = 25,36,49,64,81,100,121,144).

4.5.2. Gabor Approach

We used the method described in the paper[9] to compare the results of NMF. Gabor wavelets are localized, oriented, band pass filters. The 2-D complex valued Gabor kernel used for retrieving multi-scale and multi-orientation features[10] is given by:

$$\psi_k(\mathrm{z}) = \frac{k^2}{\sigma^2} \exp(-\frac{k^2\mathrm{z}^2}{2\sigma^2})[\exp(ik\mathrm{z}) - \exp(-\frac{\sigma^2}{2})]$$

Where value of the Gabor kernels $\psi_k(\cdot)$ is given in image coordinates and take the form of a plane wave restricted by a Gaussian envelope. The point coordinates relative to the center is given by the vector z. The parameter σ determines the ratio of window width to wavelength and the wave vector k controls the width of the Gaussian window and the wavelength and orientation of the oscillatory part. The parameters selected to construct Gabor wavelets for this experiment are as follows: three spatial frequencies (scales), $\frac{\pi}{2}, \frac{\pi}{4}, \frac{\pi}{8}$ and six levels of orientations $0, \frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}, \frac{5\pi}{6}$ were chosen respectively resulting 18 Gabor kernels. Maximum spatial frequency selected was $\frac{\pi}{2}$. The generated Gabor kernels are of size 64x64.

A face model is constructed by averaging a subset of images in the database. Nineteen feature points corresponding to landmark feature points are selected on this average face model. In selecting feature points more weightage was given to the eye area of the model as the eye area seems to be more invariant to facial expression and gestures.[13] Since eyes lie on preset coordinates of the normalized face image, we have the benefit of selecting a handful of left and right eye feature points which are aligned with this line. Three more points across the highest point of the nose were included. Left and right face contours were covered by selecting two feature points which are in line with the points selected on the nose. The remaining three points were selected on the mouth, which cover the corners of the mouth and the middle point between them.

Gabor response at each feature point is evaluated by convolving the feature point value with 18 Gabor kernels. Hence, we can represent a feature vector corresponding to a feature point as a collection of Gabor coefficients.

Feature vectors, each consisting of 18 Gabor response amplitudes are generated for each predefined feature point of the matching image and the stored images in the database. The similarity between feature vectors of two corresponding feature points on two images is determined by taking the normalized inner product of the two vectors. The similarities are used to evaluate the recognition performance.

4.5.3. Recognition Performance

We use the Receiver Operator Characteristics (ROC) to compare the efficacy of face recognition methods based on NMF, Gabor and PCA. ROC curves are used to describe the relationship between true-positive rate and false-positive rate. The corresponding thresholds are the similarities among encodings evaluated as given in Eq.(6). The identification results correspond to the likelihood of choosing the right match within the top retrieval positions. Also, we evaluate the cumulative recall rates for the n-best retrieval positions, where $n = \{1, 5, 10, 15, 20, 25\}$. Our training set $\Gamma^{train} = \{\Gamma_1{}^{train}, \Gamma_2{}^{train}\}$ consists of 834 images (male = female = 417 =(139x3)), and the test set for registered images $\Gamma^{test} = \{\Gamma_1{}^{test}, \Gamma_2{}^{test}\}$ consists of 278 images (male = female = 139x1).

As can be seen from the Figure (4)-(a)(b), the NMF technique reports around 98% true positive rate when the false positive rate is in the range of 2 and higher. Likewise, the rejection rate is comparatively higher for lower false negative rates falls within the same range. Further, we calculate the ROC power for all 3 methods, which indicates the stability and potential of each method used in the experiment. The ROC power of each technique is calculated for the false positive range 0-5% in Figure (4)-(a). The area under each curve in the range between 0-5% false positive rates in Figure (4)-(a) is evaluated and the power is calculated by taking the area of the region occupied by each ROC curve of the corresponding method. As shown in Figure (5), the results indicate that NMF has outperformed Gabor and PCA in recognition accuracy.

The likelihood of choosing the right person within the n-best (n=10) faces out of 834 images in the training set for all 3 approaches are given in Figure (6). In this case, as shown in Figure (6) the recall rate for Gabor features is higher compared to NMF at the highest precision level (n=1) and becomes almost equal when the precision level reaches the 10th n-best position. Of all methods, PCA accounted for the lowest performance with a recall rate of 40% for the 10th n-best position. In the paper by Moon,[14] PCA performance for a probe set of images taken on different dates has compared with a probe set of images taken on the same day. The results indicate a significant drop of recognition performance for images captured on different dates compare to the results achieved for the images taken on the same day. The Surrey database images were also captured on different dates similar to those in the face database used in Moon[14] and the results we achieved conform with those results.

We also bench-marked the behavior of parts-based NMF, when imposters are introduced for recognition. For this, we use the test set γ against the training set Γ^{train}. Of 17 imposters, only one was falsely accepted as a match in the database supporting the fact that likelihood of retrieving a close match for an unseen image is low compared to the retrieving of a registered image.

4.6. Experiment 3: Image Reconstruction with gender-based features

The goal of this experiment is to explore the behavior of gender-based NMF features during the face reconstruction. Often, more abstract features like gender are determined from local templates which use grey values corresponding to image pixels[15].[16] We seek to explore the influence of low-level NMF parts-based features in discriminating genders based on the reconstruction error. The error function given in Eq.(7) is used to calculate the quadratic reconstruction error for the experimental image sets where the value of ε is considered as an inverse measure of certainty for a given gender class. When an image is reconstructed with basis images from one or the other gender class, the basis set of the synthesized image with the lowest reconstruction error determines the class membership of the target image.

The following experiment is carried out to analyze the behavior of NMF features based on the gender type. The female test ensemble $\Gamma_1{}^{test}$ is reconstructed using a set of basis images, learned entirely on female images $\Gamma_1{}^{train}$ and a set of basis images, learned entirely on male images $\Gamma_2{}^{train}$. Of 139 female images in the $\Gamma_1{}^{test}$ set, 137 images showed lower reconstruction error when reconstructed with female basis images than with the male basis images. This suggests that the set of features extracted from male images shows important differences from
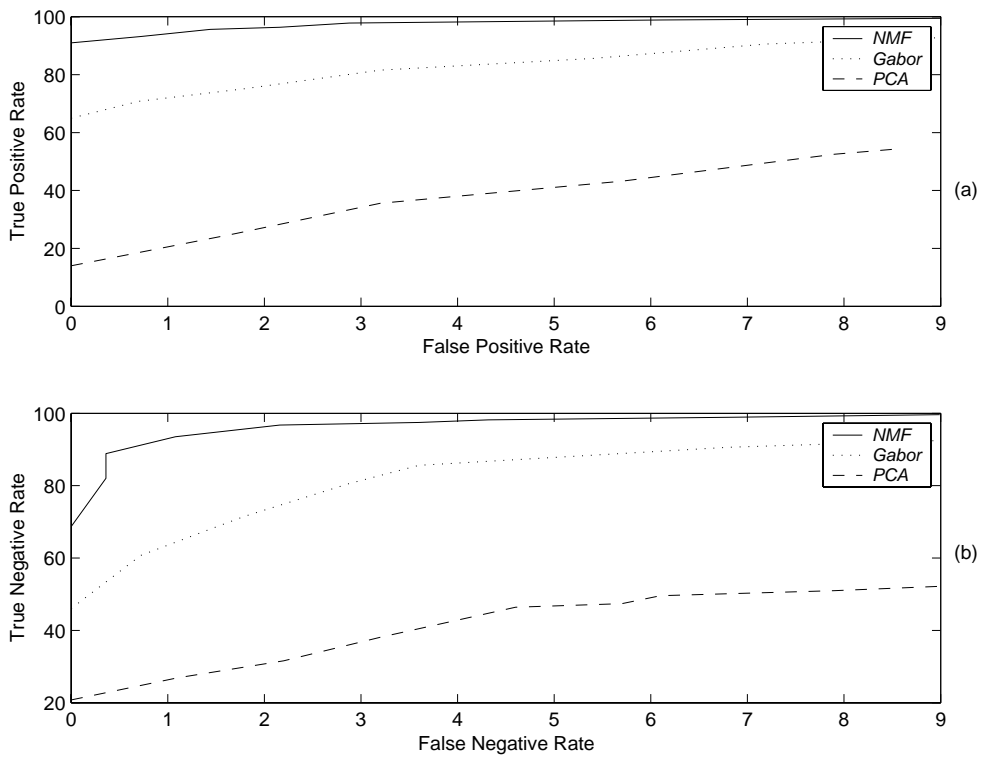
Figure 4. (a)True positive rate(%) vs False positive rate(%) (b) True negative rate vs False negative rate; where thresholds are the similarities between the encodings of test and train faces.
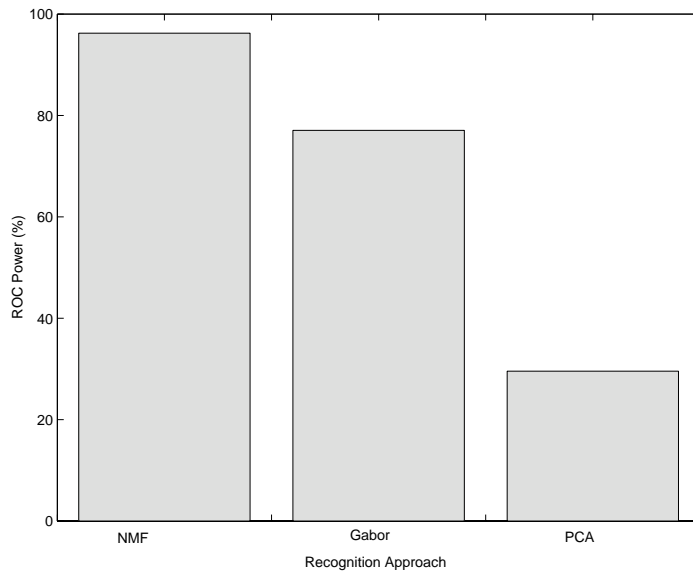

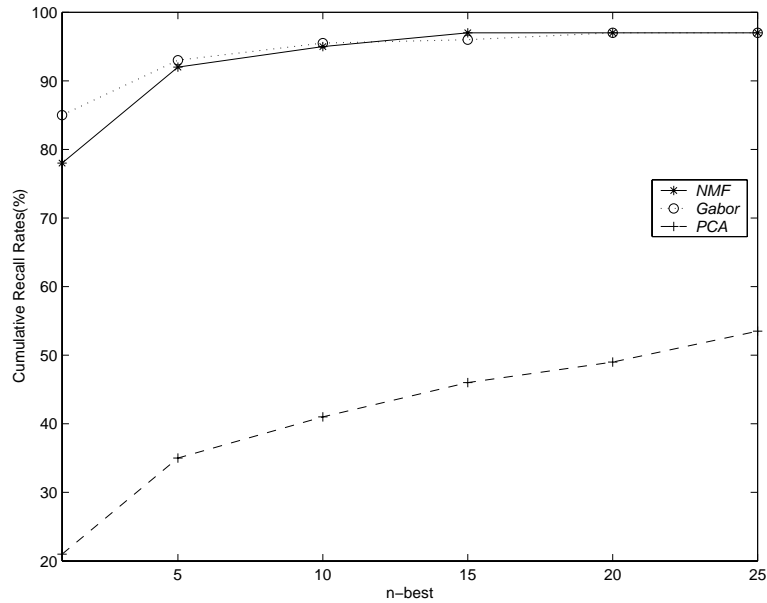**Figure 5.** ROC power (%) for NMF, Gabor and PCA.

Figure 6. Cumulative Recall Rates (%) for NMF, Gabor and PCA for n-best (n=1,5,10,15,20,25) for a set of 278 images of equal number of female and male images.

the feature set extracted from female images. There may be features exclusive to female images which do not appear in male images. Likewise, we reconstructed the male test ensemble Γ_2^{test} with all female and all male basis images separately. Again, of 139 male images, 129 images gave low reconstruction error when reconstructed with male basis images, compared to the ones reconstructed with female basis images. For the male image set however, the performance was not as good as the female set with 10 male images reconstructed with female basis images resulted in lower reconstruction errors. This suggests that if male feature space may be a subset of female feature space with certain specific features contributing most in establishing gender orientation.

We hypothesize that there exist unique features or feature combinations (or encodings) that influence the accurate discrimination of genders. To further explore this hypothesis, we used an imposter set γ, and test the gender of these previously unseen images based on the lowest reconstruction error. The basis images and encodings were derived from both training sets Γ_1^{train} and Γ_2^{train}, and were used separately to reconstruct the imposter set. Of 68 male imposters, 62 were correctly categorized as male images indicating the capability of low level features to contribute to gender determination. Having faces of different races in the database can also be a possible factor for the wrong categorization of some of the images. It has been shown that gender specific features are not completely constant across races of faces.[17] Better results can be envisioned by combining geometrical features together with parts-based local features.

5. CONCLUSIONS

We have applied parts-based NMF to learn face images from the M2VTS database. Experimental results show that this method is quite robust and yields better results compared to described Gabor feature approach and PCA for face recognition. The recognition performance we achieved for PCA is comparatively low, but conforms with the results achieved in literature for the images taken over a period of time.

Even though the results we achieved in attempting to discriminate genders using local parts-based features alone didn't yield high discrimination performance, these results infer that these features contribute to the process of gender discrimination. The NMF-based features seem to facilitate the reconstruction of images of the correct gender with a lower reconstruction error at a rate of 91% indicating the suitability of parts-based features in

gender determination. The wrong gender classification may have caused by having faces of different races in the database.

In our experiment, an equal number of male and female images when reconstructed with own and opposite gender basis images, more female images were able to reconstruct with lower reconstruction error with female basis images compared to its male counter part reconstructed with male basis images. This posed the question whether male features are just a subset of female feature space representing only a fewer number of features than female features, which are sufficient to represent male faces.

Natural images such as face images that can be described as a collection of a small number of landmark components seem to lend themselves to parts-based learning and demonstrates NMF's suitability for face recognition applications. Even though face recognition algorithms yield better results with local feature based methods due to the robustness to variabilities in the local regions, we can not ignore the influence of global properties towards the final recognition accuracy. Often, we recognize people seen before from a distance implying that we do not tend to formalize a picture of that person in our visual system based on the facial parts alone. Other factors such as hair, gait and physique seem to play a big role in overall identification process.

REFERENCES

1. M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cognitive Neuroscience* **3(1)**, pp. 71–86, 1991.
2. P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces:recognition using class specific linear projection," **19(7)**, pp. 711–720, 1997.
3. D. J. Field, "What is the goal of sensory coding?," *Neural Computation* **6**, pp. 559–601, 1994.
4. B. A. Olshausen and D. J. Field, "Sparse coding with an overcomplete basis set: A strategy employed by v1?," *Vision Research* **37**, pp. 3311–3325, 1997.
5. D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature* **401**, pp. 788–791, 1999.
6. D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," *Proc. NIPS*, pp. 556–562, 2000.
7. B. S. Manjunath and et al., "A feature based approach to face recognition," *IEEE Computer Vision and Pattern Recognition*, pp. 373–378, 1992.
8. L. Wiskott and et al., "Face recognition by elastic bunch graph matching," *IEEE International Conference on Image Processing* **1**, pp. 129–132, 1997.
9. M. Rajapakse and Y. Guo, "Efficient gabor feature based approach for face recognition," *Proc. World Multi Conference on Systemics,Cybernetics and Informatics* **5(1)**, pp. 265–269, 2000.
10. J. G. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *J. of Optical Society of America* **2(7)**, 1985.
11. M. P. Homewood, "Dual perspectives give science added insight into brain," *The Gazette Online, The Newspaper of the Johns Hopkins University* **32(2)**, 2002.
12. X. Xiangrong and et al., "Learning representative local features for face recognition," *IEEE CVPR*, 2001.
13. Z. Zhang, "Feature-based facial expression recognition: sensitivity analysis and experiments with a multi-layer perceptron," *International Journal of Pattern and Artificial Intelligence* **13(6)**, pp. 893–911, 1999.
14. H. Moon and et al., "Analysis of pca-based face recognition algorithms," in *Empirical Evaluation Techniques in Computer Vision*, K. W. Boyer and P. J. Phillips, eds., IEEE Computer Society Press, 1998.
15. B. Moghaddam and et al., "Gender classification with support vector machines," *Proc. of Int'l Conf. on Automatic Face and Gesture Recognition*, pp. 306–311, 2000.
16. B. Edleman, "Sex classification of face areas: how well can a linear neural network predict human performance," *J. of Biological System* **6(3)**, pp. 241–264, 1998.
17. A.J.O'Toole and et al., "An other-race effect for categorizing faces by sex," **25**, pp. 669–676, 1996.