行政院國家科學委員會專題研究計畫 期中進度報告

多媒體影音高階處理、傳輸及設計--子計畫一:球賽視訊的場景偵測，分類與總覽(2/3)

期中進度報告(精簡版)

計 畫 類 別 ：整合型
計 畫 編 號 ：NSC 95-2221-E-002-189-
執 行 期 間 ：95 年 08 月 01 日至 96 年 07 月 31 日
執 行 單 位 ：國立臺灣大學電機工程學系暨研究所

計 畫 主 持 人 ：貝蘇章

處 理 方 式 ：期中報告不提供公開查詢

中 華 民 國 96 年 06 月 14 日

球賽視訊的場景偵測，分類與總覽—子計畫(2/3)
Semantic Scenes Detection, Classification, and Summarization in Sports Video(II)

計畫編號: NSC 95-2221-E-002-189

執行期限：95 年 8 月 1 日至 96 年 7 月 31 日

主持人：貝蘇章　　台灣大學電機系教授

摘要

本論文的研究主題是智能型視訊及影像分析，人臉偵測及辨識將是重要課題，利用以基底為基礎的非負值矩陣的拆解，進行人臉辨識，其效果良好，實驗結果及與其他方法的比較將在論文中展示。

ABSTRACT

A fundamental problem of Non-negative Matrix Factorization (NMF) is that it does not always extract basis components manifesting localized features which are essential in face recognition. The aim of our work is to strengthen localized features in basis images and to impose orthonormal characteristic of Principle Component Analysis (PCA) on NMF. Such improved technique is called Basis-emphasized Non-negative Matrix Factorization (BNMF). In order to reduce noise disturbance in the original image such as facial expression, illumination variation and partial occlusion, Wavelet Transform (WT) is applied before the BNMF decomposition. In this paper, a novel subspace projection technique, called Basis-emphasized Non-negative Matrix Factorization with Wavelet Transform (wBNMF), is proposed to represent human facial image in low frequency sub-band and yields better recognition accuracy. These results are compared with those produced by PCA and NMF.

1. INTRODUCTION

Face recognition is one of the most challenging problems to be solved in the computer vision community due to the wide variety of illumination condition, facial expression and occlusion. It has several potential applications in areas such as Human Computer Interaction (HCI), biometrics and security. Moreover, it is a prototypical pattern recognition problem whose solution is helpful in many other classification problems.

Until now, several sophisticated approaches have been developed to obtain better recognition result using some face databases. But there is no uniform way to establish the best approach because nearly all of them have been designed to work with faces under some specific situations. In order to obtain comparable result, the three databases used are the MIT CBCL face database [1], the Cambridge ORL face database [2] and the Aleix Robert (AR) face database [3].

One effective approach for face recognition is Principal Component Analysis (PCA) [4] which can simplify a dataset by transforming the data to a new coordinate system with the greatest variance. PCA learns basis components for subspace representation and achieves dimension reduction by discarding the least significant components. The eigenimage method uses PCA [5] [6] performed on a set of training images to decorrelate second order moments corresponding to low frequency property. Each input image can be represented as a linear combination of these eigenimages. Due to the holistic nature of this method, the resulting components are global interpretations, and thus PCA is unable to extract basis components manifesting localized features. However, localized features offer advantages in object recognition, including stability to local deformation, lighting variation, and partial occlusion. Therefore, several methods have been proposed for localized, part-based feature extraction.

Recently a subspace method called Non-negative Matrix Factorization (NMF) is proposed by Lee and Seung [7] [8] as a way to find a set of basis functions for representing non-negative data, which has been used for image representation, document analysis [7] and clustering [9] [10] [11] for its parts-based representation property. NMF is akin to other matrix decompositions which have been proposed previously, such as Positive Matrix Factorization (PMF) of Juvela, Lehtinen and Paatero [12] [13]. The non-negativity constraints make the representation purely additive, in contrast to many other linear representations such as PCA and Independent Component Analysis (ICA) [14] [15]. However, the additive parts learned by NMF are not necessarily localized for some databases such as the ORL face database. Experiments also show that directly using the learned feature vectors via NMF under the Euclidean distance cannot get better face recognition rate than that obtained by the traditional PCA. In order to improve the recognition accuracy, Local Non-negative Matrix Factorization (LNMF) [16] is proposed to achieve a more localized NMF algorithm with the aim of computing spatially localized bases from a face database by adding three constraints that modify the objective function in the original NMF algorithm. But this method has a slow speed for learning the bases. Then, Guillamet and Vitrià adopt one relevant metric called Earth Mover's Distance (EMD) [17] for parts-based representation of NMF. However, the computation of EMD is too time-demanding. Recently Hoyer incorporated the notion of sparseness to improve the found decompositions, and then proposed a method called Non-Negative Matrix Factorization with Sparseness Constraints (NMFSC) [20] [21]. But its recognition accuracy is not better than that of PCA.

In this paper, a novel subspace method is proposed, called Basis-emphasized Non-negative Matrix Factorization with Wavelet Transform (wBNMF), for learning intuitive parts-based representation of visual pattern with noise reduction [22] [23]. Inspired by previous work, our aim is to impose orthonormal characteristic on basis components and to make the representation more suitable for tasks where feature localization is important. This paper also investigates how to improve the face recognition accuracy based on wBNMF [24] [25] [26] [27] [28] [29]. For better performance, we adopt the Riemanian metric distance for the learned feature vectors instead of the Euclidean distance [18] [19]. Experiments on the widely used AR face database demonstrate the proposed method can improve recognition accuracy and even outperform PCA.

2. FACE DATABASE

2.1. CBCL face database

There are 2429 faces and 4548 non-faces in the training set. The testing set consists of 472 faces, 23573 non-faces. Each image is 19×19 grayscale PGM format. Figure 1 shows some sample images from the database.

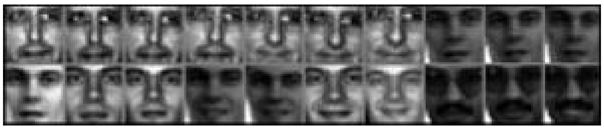


Figure 1: Face examples from the CBCL database.

2.2. ORL face database

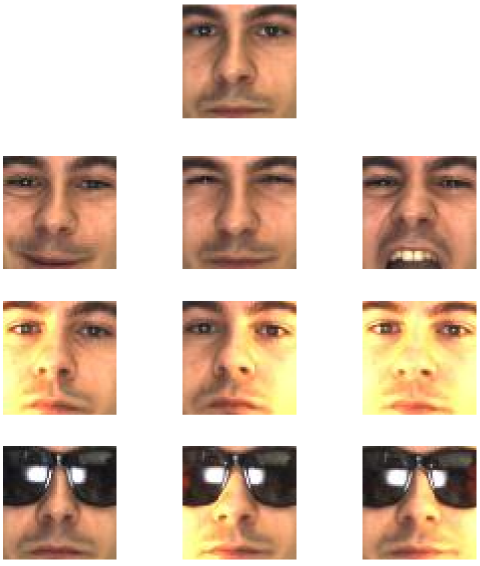
There are 400 face images of 40 persons, 10 images per person which are shown in Fig. 2. The original dimension of each image is 112×92. These images are taken at different times, slightly varying lighting, facial expressions (open/closed eyes, smiling/non-smiling) and facial details (glasses/no-glasses). All the images are taken against a dark homogeneous background. The faces are in up-right position of frontal view, with slight left-right out-of-plane rotation. Each image is linearly stretched to the full range of pixel value within [0, 255]. In Fig. 3, the ORL faces are re-aligned in the center of the original images. The redundancy region of each image, non-facial feature, is eliminated to avoid undesired noise in face image analysis. The dimension of the normalized ORL face image is 48×48.

Figure 2: Face examples from the ORL database.

Figure 3: Face examples from the normalized ORL database.

2.3. AR face database

The AR color face database contains images of 126 individuals (70 males and 56 females). Original images are 768×576 pixels in size with 24-bit color resolution. A total of 13 photos are taken from each individual with each shot taken under different conditions as shown in Fie. 4. These same shots are taken again after two weeks interval in another session. For our experiments, only 200 face images (50 males and 50 females) of 13 shots in both sessions of the original AR database were randomly extracted. In order to avoid external influence of background, these realigned images are cropped and down-sampled in such a way that the final image size is 120×120 pixels.



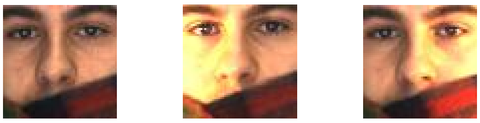


Figure 4: Face examples from the cropped AR database (from left to right, from top to down) : neutral, smile, anger, scream, left light, right light, both lights, sunglasses, sunglasses & left light, sunglasses & right light, scarf, scarf & left light, scarf & right light.

3. BASIS-EMPHASIZED NON-NEGATIVE MATRIX FACTORIZATION

3.1. Drawbacks of NMF

One noticeable property of NMF is that it usually produces a naturally sparse representation of the input data. Such a representation encodes much of the input data by using relatively few active bases, which make the encoding easy to interpret. Lee and Seung (1999) originally showed that NMF learned a parts-based representation when trained on the CBCL database.
Despite this success, when applied to the ORL database in which images are not aligned as well, a global decomposition emerges. Therefore, the difference in result was mainly attributed to how well the images were hand-aligned (Li et al., 2001). In Fig. 5, NMF is applied to different face databases, the CBCL and ORL databases. The representation of basis images learned from the CBCL database is apparently composed of the intuitive facial features, but the representation of basis images learned from the ORL database is global rather than local.

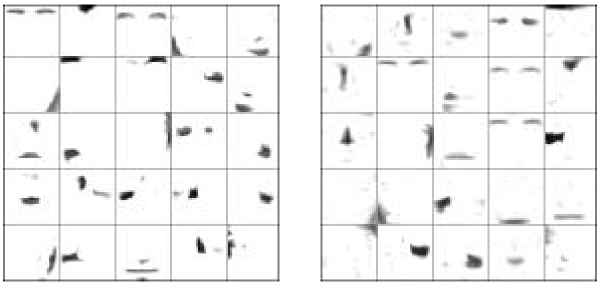
(a) (b)

Figure 5: Basis images learned from the CBCL (a) and ORL (b) database using NMF.

3.2. Extensions of NMF

Sparseness in both the bases and encodings is crucial for a parts-based representation. For this reason, many studies incorporating the notion of sparseness are developed to improve the found decomposition. One useful approach is non-negative matrix factorization with sparseness constraints (NMFSC), the aim of which is to constrain NMF to find decomposition with desired degree of sparseness.

In the other point of view, the manifestation of localized features is significant and then an improved method, local non-negative matrix factorization (LNMF), is proposed. LNMF defines a novel objective function with additional localization constraint. However, many basis images learned from the normalized ORL database using LNMF and NMFSC respectively lack intuitive facial features as shown in Fig. 6. Moreover, most of them are just non-meaningful fragments. In addition, the sparseness of the basis images in NMFSC is fixed at 0.75 and higher pixel values are in darker color. In the following cases, the normalized ORL face image database is chosen to avoid undesired noise.

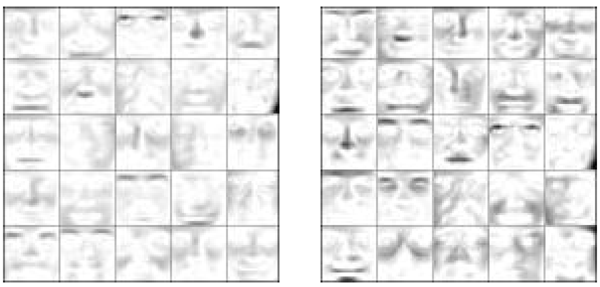


(a) (b)

Figure 6: Basis images learned from the normalized ORL database using LNMF (a) and NMFSC (b) respectively.

3.3. Algorithm of BNMF

The basis images we desire are non-global and contain several versions of mouths, noses and other intuitive facial features in different locations and forms. In order to meet human intuitive notion of an individual face, some symmetrical facial features have better come out in pairs such as eyes and eyebrows. These considerations lead us to designing an improved NMF which can learn basis images with weakened holistic contour and emphasized local features. The accomplished result is shown in Fig. 7 and it is noticeable that our method can learn desired facial features, in contrast to the original NMF.



(a) (b)

Figure 7: Basis images learned from the normalized ORL database using NMF (a) and BNMF (b).

Since the original NMF does not impose any constraints on the spatial localization, minimizing the objective function hardly yields a factorization which can reveal local features in the basis images. Therefore, BNMF is introduced to impose more spatial constraints on the cost function. A new objective function is defined to learn intuitively parts-based components. Letting $B=[b_{ij}]=W^T W$, BNMF can learn local features by imposing the following constraints on the bases and encodings.

(1) Given the existing constraint $\sum_i w_{ij}=1$ for all i, we wish that each basis should be as orthogonal as possible, so as to minimize redundancy between different bases. The objective function can be imposed by $\sum b_{ij}=\min$

(2) Take advantage of essential characteristics in PCA, we make each basis image have constant energy. This can be achieved by $W_{ik} \leftarrow \frac{W_{ik}}{\sqrt{\sum_{p=1}^{n} W_{pk}^2}}$

(3) To avoid the degradation of energy on coefficients, we renormalize each row of H to maintain constant energy. This can be done by $H_{kj} \leftarrow \frac{H_{kj}}{\sqrt{\sum_{p=1}^{m} H_{kp}^2}}$

The incorporation of the above constraints leads to the following objective function for BNMF:

$$F = \sum_{i=1}^{n}\sum_{j=1}^{m}(V_{ij}-(WH)_{ij})^2+\sum B_{ij}$$

When subject to the non-negativity for all matrix factors, a local solution to the minimization of such a constrained objective function can be found under the update rules:

$$W_{ik} \leftarrow W_{ik}\sum_{j=1}^{m}\frac{V_{ij}}{(WH)_{ij}}H_{kj}$$

$$W_{ik} \leftarrow \frac{W_{ik}}{\sum_{p=1}^{n} W_{pk}}$$

$$W_{ik} \leftarrow \frac{W_{ik}}{\sqrt{\sum_{p=1}^{n} W_{pk}^2}}$$

$$H_{kj} \leftarrow H_{kj}\sum_{i=1}^{n} W_{ik}\frac{V_{ij}}{(WH)_{ij}}$$

$$H_{kj} \leftarrow \frac{H_{kj}}{\sqrt{\sum_{p=1}^{m} H_{kp}^2}}$$

3.4. Comparison of NMF-related Algorithms

BNMF adds the non-negative constraint and the orthonormal characteristic of PCA to get intuitive parts-based features. Therefore, bases should be as orthogonal as possible so as to minimize redundancy between them.

To show the advantage of BNMF, we compare it with other extensions of NMF. The Euclidean distance and the divergence distance between the original and reconstructed images are computed to evaluate the efficiency of each algorithm. The smaller distance value we compute, the better matrix factorization we get. Furthermore, the influence of various dimensions (number of basis components) on the result basis images is surveyed.

As Fig. 7 shows, NMF and BNMF learn basis images which contain dark intuitive part-based facial features and light global facial contour. Higher contrast between the holistic contour and the local feature emerges in BNMF. Nevertheless, some of the basis images learned from LNMF and NMFSC in Fig. 6 are no more than non-meaningful fragments. In Table 1, we intentionally adopt two different kinds of distance metrics, the Euclidean distance and the divergence distance. So we can see more deeply that the smallest distance value occurs in BNMF.

Table 1: Distance between the original and the reconstructed images under various extensions of NMF and NMF itself. And we set the dimension for 25 and the iteration time for 1000.

Algorithm	Euclidean distance	Divergence distance
NMF	$5.2398e+003$	$5.5763e+003$
LNMF	$2.9566e+005$	$6.6422e+005$
NMFSC	$1.9498e+004$	$1.9731e+004$
BNMF	$5.1654e+003$	$5.4661e+003$

4. WAVELET TRANSFORM

4.1. Two-Dimensional Discrete Wavelet Transform

Since the information the continuous wavelet transform (CWT) provides is highly redundant as far as the reconstruction of the signal is concerned, the discrete wavelet transform (DWT) is needed to provide sufficient information both for analysis and synthesis of the original signal, with a significant reduction in the computation time. DWT employs two sets of functions, called scaling functions and wavelet functions obtained by successive high pass and low pass filtering of the time domain signal. The equation of DWT is given:

$$(W_y\, f(x))(m,n) = \frac{1}{\sqrt{a_0^m}}\sum_k x[k]y\,[a_0^{-m}n-k]$$

Where *m* is a scaling integer variable, *n* is a shifting integer variable, *x[k]* is a digital signal with sample index *k*, and y is the mother wavelet.

For images, two-dimensional DWT is needed to decompose approximation coefficients at level *j-1* to four components at level j which are the approximation (LL_j) and the details in three orientations: horizontal (LH_j), vertical (HL_j) and diagonal (HH_j). On the other hand, two-dimensional discrete wavelet transform (IDWT) is used to reconstruct the original image. Fig. 8 describes the basic decomposition and reconstruction steps.

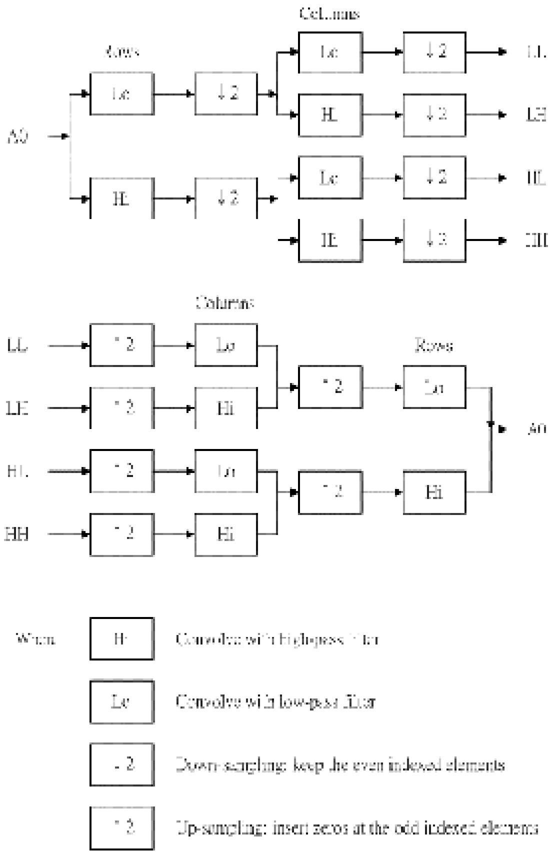
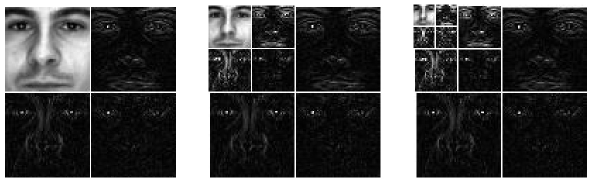


Figure 8: The decomposition and reconstruction of two-dimensional DWT.

4.2. Wavelet Sub-Bands

Two-dimensional discrete wavelet transform can be used to decompose the facial images into a multi-resolution representation in order to keep the least coefficients possible without losing useful image information. Fig. 9 depicts the decomposition process by applying the two-dimensional Haar wavelet transform of a face image and depicts the successive levels wavelet decomposition by applying the Haar wavelet transform on the low-frequency band sequentially. Note that the highest-frequency wavelet sub-band contains mostly noise and the contour of the decomposed facial image is clearer toward the left-top direction.



(a) (b) (c)

Figure 9: Face image in wavelet sub-bands. (a) 1-level wavelet decomposition, (b) 2-level wavelet decomposit-ion, (c) 3-level wavelet decomposition.

In this paper, Wavelet Transform is used to reconstruct a better representation in the low spatial frequency bands by discarding the highest frequency spectrum of each level as shown in Fig. 10(d). Hence, it can make the facial images insensitive to facial expression, illumination variation and occlusion. During reconstruction, these discarded coefficients are replaced with zeros. We also compute the difference between the reconstructed and the original images and the error result is shown in Fig. 11, in which the larger the error, the lighter the color for each element.

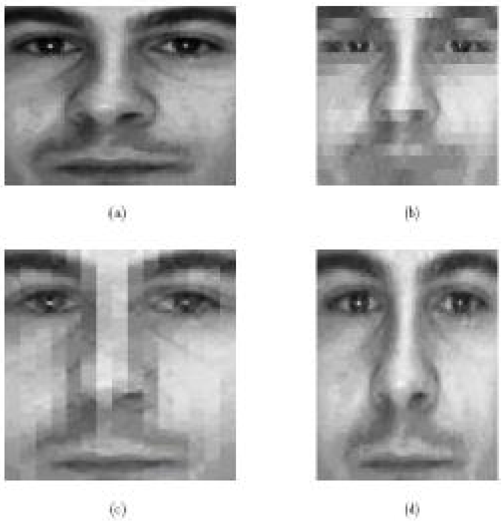


Fig 10: The reconstructed images using two-dimensional IDWT discarding (a) no frequency band, (b) the first (lowest) frequency band, (c) the second frequency band, (d) the third (highest) frequency band.

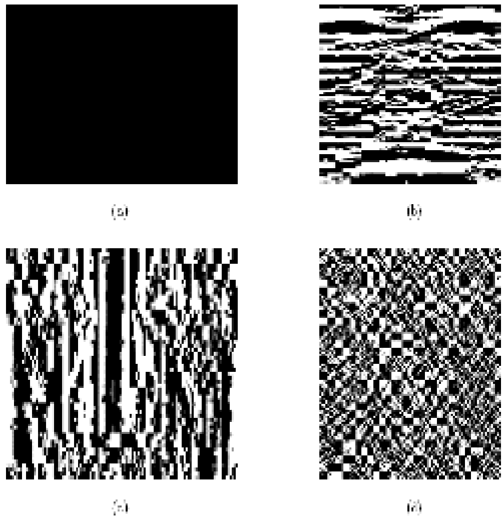


Figure 11: The error images between the original image and the corresponding reconstructed image in Fig. 10.

4.3. Basis-emphasized Non-negative Matrix Factorization with Wavelet Transform

This section introduces a novel subspace projection technique via BNMF to represent human facial image in low frequency sub-band, which is able to realize through the Wavelet Transform. After wavelet decomposition and reconstruction, BNMF is performed to produce part-based representations of the images. The simulation results on the AR database show that the hybrid of BNMF and the best wavelet filter will yield better recognition rate and shorter training time. In order to achieve an excellent verification rate when identifying the faces, We investigate the performance obtained by the integration of WT and BNMF to take the advantages of these two methods. These results are compared with those learned by PCA and the original NMF techniques later.

In face recognition, dimensionality reduction is very important to project the facial images from a high-dimensional space onto a lower-dimensional space. And wavelet transform reduces the resolution of images and decreases the computation load of feature generation. With the adoption of wavelet transform, the training time can also be reduced significantly. In this paper, three level of wavelet decomposition is performed on face images. The reconstructed face image discarding the highest-frequency sub-band is then subjected to BNMF. The integrated framework of Wavelet Transform and Basis-emphasized Non-negative Matrix Factorization is abbreviated as wBNMF. The flow chart of wBNMF is illustrated in Fig. 12.

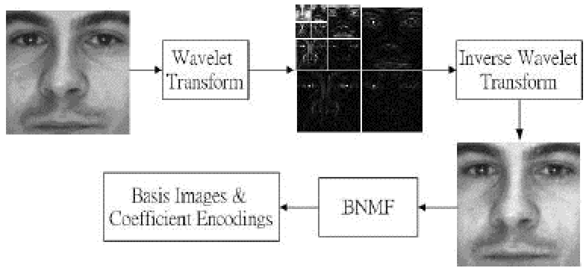


Figure 12: Flow chart of generating the wBNMF features.

5. FACE RECOGNITION

5.1. Classifier Determination

In this experiment, we adopt the Riemannian distance metric which is more suitable than any other distance metric for face classification when using the nearest neighbor classifier. Let f_1, f_2 denote two facial vectors in the original n-dimension space, and the corresponding learned coefficients in lower r-dimension space are h_1, h_2, respectively. To some extent, we can say $f_1 = Wh_1$ and $f_2 = Wh_2$ where W is the learned basis matrix. Then we get

$$Rie(f_1, f_2) = (f_1 - f_2)^T (f_1 - f_2) = (h_1 - h_2)^T W^T W (h_1 - h_2)$$
$$= (h_1 - h_2)^T G (h_1 - h_2) \neq (h_1 - h_2)^T (h_1 - h_2)$$

This indicates that the Riemannian metric can preserve the neighborhood of the original samples for classification. In addition, the technique is able to improve recognition accuracy when a higher rank is used.

5.2 Facial Expression

In order to see how each technique can deal with expression, facial images in the AR face database are used as a testing set because they contain smile, anger and scream expressions. And we use the neutral facial expression images as the training set. NMF and wBNMF techniques are executed under 1000 iteration time.

Given scream facial images as a testing set, it is noticeable that PCA produces the poorest recognition rate. Since PCA is based on learning holistic nature, it mostly extracts global features of the original images and cannot handle obvious distortion such as scream facial expressions. On the contrary, the other two parts-based methods, NMF and wBNMF, are more suitable to solve such a problem, especially with the significant performance improvement on behalf of wBNMF under the lowest feature dimensional space. And wBNMF outperforms both PCA and NMF for smile and anger expressions under such feature dimensional space. Therefore, the small basis number is necessary for wBNMF to reach good recognition accuracy. The reason is that smaller basis number is helpful to produce more robust wBNMF basis representation which tends to neglect the seriously distorted facial part.

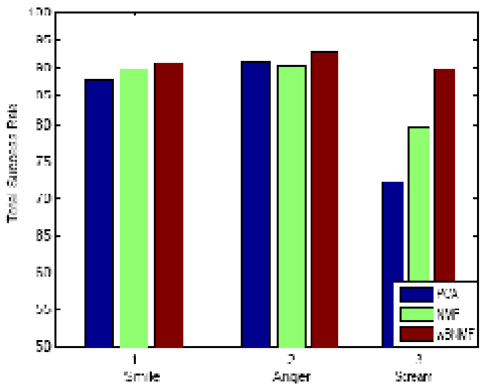


Figure 13: Total success rate versus various facial expressions using techniques, PCA, NMF and wBNMF.

Fig. 13 shows the comparison of recognition performance when using different facial expressions as testing sets. All facial expressions are relatively better classified when using wBNMF under the lowest feature dimensional space. But scream faces demonstrate the worst verification rate because of its striking facial expression where PCA and NMF are unable to deal with.

Finally, we conclude that the wBNMF technique is more suitable than the other two to recognize various facial expressions.

5.3. Illumination Variation

Illumination condition is a factor which should be taken into account in face recognition. The condition is reflected in neutral facial images under three kinds of lighting way: left light on, right light on and both lights on. We experimentally demonstrate that illumination condition is a serious matter in face recognition. As what we anticipate, harsher illumination change reduces recognition accuracy.

Under varying illumination conditions and constant facial expression, the experimental result shows that PCA and NMF can not deal with illumination variations as good as wBNMF. Due to the noise reduction caused by Wavelet Transform, wBNMF is relatively good to withstand global changes of image. And it can improve the ability of NMF to handle illumination variations. Therefore, wBNMF deals with illumination variations a little better than the other two, PCA and NMF. In addition, when the lighting scope is expanded from half to whole facial region, the recognition rate decreases obviously about thirty percent as shown in Fig. 14.

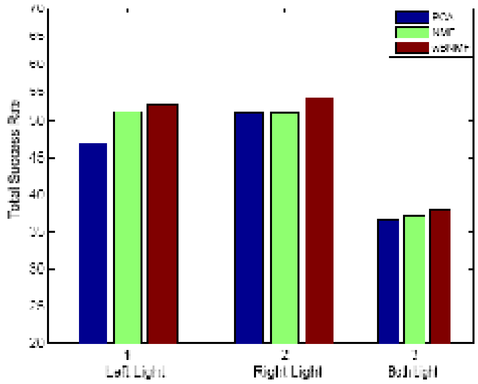


Figure 14: Total success rate versus various lighting conditions using techniques, PCA, NMF and wBNMF.

5.4. Occlusion Disturbance

Here we have a set of natural occlusion where faces are occluded with a scarf or sunglasses. It means that the eyes and mouth in the facial image are occluded. Under the presence of sunglasses or scarf, recognition rates decrease considerably. This implies that certain facial features, such as eyes and mouth, are very important for classifying faces as shown in Fig. 15.

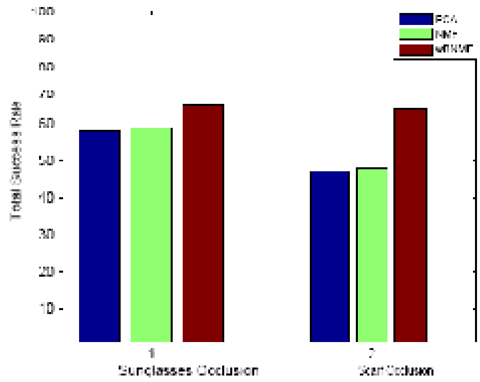


Figure 15: Total success rate versus facial images with sunglasses or scarf occlusion using techniques, PCA, NMF and wBNMF.

The result clearly shows that wBNMF outperforms the other approaches by a large margin when the occluded region, such as eyes and mouth, is significantly large and crucial for face recognition. On the contrary, PCA can not manage such conditions as well because it is focused on extracting global face features. For this reason, wBNMF is able to classify occluded facial images with the best recognition rate, even comparable to the best one obtained by PCA or NMF. As a result, we believe that wBNMF can be a relevant technique for pattern recognition problems, where partial face occlusion that can not be handled by PCA may appear.

For a specific pattern such as human face, the recognizable facial features occupy only a fraction of all.

Since a specific pattern of interest can reside in a low dimensional sub-manifold of the original input data space which consists of an unnecessarily high dimensionality. One of subspace analyses, wBNMF, is used to reveal low dimensional structures observed in a high dimensional space. In fact, the essence of feature extraction in pattern recognition can be considered as discovering and computing low dimensional intrinsic pattern from observation. Subspace analysis has demonstrated its success in numerous visual recognition tasks such as face recognition and detection.

6. REFERENCES

[1] MIT Center for Biological and Computation Learning, "The CBCL face database," Available at: http://cbcl.mit.edu/cbcl/software-databasets/FaceData2.html

[2] AT & T Laboratories, Cambridge, U.K., "The ORL face database," Available at: http://www.orl.co.uk/facedatabase.html

[3] A. M. Martinez and R. Benavente, "The AR face database," *Technical Report 24, Computer Vision Center (CVC)*, Barcelona, Spain, June 1998.

[4] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 586-591, Hawaii, June 1991.

[5] L. Sirovich and M. Kirby, "Low-dimensional procedure for the characterization of human faces," *Journal of the Optical Society of America*, Vol. 4, No. 3, pp. 519-524, March 1987.

[6] M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve procedure for the characterization of human faces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 1, pp. 103-108, January 1990.

[7] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, pp. 788-791, 1999.

[8] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," *Advances in Neutral Information Processing 13 (Proc. NIPS*2000)*, MIP Press, 2001.

[9] D. Guillamet and J. Vitriȁ, "Unsupervised learning of part-based representation," *Computer Analysis of Images and Patterns, W. Skarbek (Eds.) Lecture Notes in Computer Science 2124*, Springer, pp. 700-708, 2001.

[10] Wei Xu, Win Liu, and Yihong Gong, "Document clustering based on non-negative matrix factorization," *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval*, July 28-August 01, 2003, Toronto, Canada.

[11] Jeffrey Ho, Ming-Hsuan Yang, Jongwoo Lim, Kuang-Chih Lee, and David Kriegman, "Clustering Appearances of Objects Under Varying Illumination Conditions," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 11-18, 2003.

[12] M. Juvela, K. Lehtinen, and P. Paatero, "The use of positive matrix factorization in the analysis of molecular line spectra from the thumbprint nebula," In D. P. Clemens and R. Barvainis, editors, *Clouds, Cores, and Low Mass Stars, ASP Conference Series*, vol. 65, pp. 176-180, 1994.

[13] M. Juvela, K. Lehtinen, and P. Paatero, "The use of positive matrix factorization in the analysis of molecular line spectra," *MNRAS*, vol. 280, pp. 616-626, 1996.

[14] P. Comon, "Independent component analysis -- a new concept?," *Signal Procesing*, vol. 36, pp. 287-314, 1994.

[15] Hyvärinen, J. Karhunen, and E. Oja, "Independent Component Analysis," *Wiley Interscience*, 2001.

[16] S. Z. Li, X. Hou, H. Zhang, and Q. Cheng, "Learning spatially localized parts-based representations," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. I, pp. 207-212, Hawaii, USA, 2001.

[17] D. Guillamet and J. Vitriȁ, "Evaluation of distance metrics for recognition based on non-negative matrix factorization," *Pattern Recognition Letters*, vol. 24, pp. 1599-1605, June 2003.

[18] D. Guillamet and J. Vitriȁ, "Determining a suitable metric when using non-negative matrix factorization," *Proceedings of the 16th International Conference on Pattern Recognition (ICPR 2002)*, pp. 128-131, Quebec, 2002.

[19] W. Liu and N. Zheng, "Non-negative matrix factorization based methods for object recognition," *Pattern Recognition Letters*, vol. 25, no. 8, pp. 893-897, June, 2004.

[20] P. O.~Hoyer, "Non-negative sparse coding," *Neural Networks for Signal Processing XII (Proc. IEEE Workshop on Neural Networks for Signal Processing)*, pp. 895-898, The MIT Press, Cambridge, Massachusetts, 1995.

[21] P. O. Hoyer, "Non-negative matrix factorization with sparseness constraints," *Journal of Machine Learning Research*, vol. 5, pp. 1457-1469, November 2004.

[22] D. Donono and V. Stodden, "When does non-negative matrix factorization give a correct decomposition into parts?" *Proceedings of NIPS workshop on ICA*, 2003.

[23] N. H. Foon, A. T. B. Jin, and D. N. C. Ling, "Face recognition using wavelet transform and non-negative matrix factorization." *Australian Conference on Artificial Intelligence*, pp. 192-202, 2004.

[24] D. Guillamet and J. Vitriȁ, "Non-negative matrix factorization for face recognition," *Escrig Monferrer, M. T., Toledo, F., Golobardes, E. (Eds.), Topics in Artificial Intelligence. Springer-Verlag Series: Lecture Notes in Artificial Intelligence*, vol. 2504, pp. 336-344, Castello de la Plana, 2002.

[25] D. Guillamet and J. Vitrià, "Classifying faces with non-negative matrix factorization," *Proceedings of the 5th Catalan Conference for Artificial Intelligence (CCIA 2002)*, pp. 24-31, Castello de la Plana, 2002.

[26] Xiaorong Pu, Zhang Yi, Ziming Zheng, Wei Zhou, and Mao Ye, "Face recognition using fisher non-negative matrix factorization with sparseness constraints," *Advances in Neural Networks - ISNN 2005: Second International Symposium on Neural Networks*, Chongqing, China, May 30 - June 1, 2005, *Proceedings, Part II Lecturer notes in computer science*, vol. 3497, pp. 112-117, April 2005.

[27] Y. Wang, Y. Jia, C. Hu, and M. Turk, "Non-negative matrix factorization frame work for face recognition," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 19, no. 4, pp. 1-17, 2005.

[28] I. Buciu and I. Pitas, "A new sparse image representation algorithm applied to facial expression recognition," *Proceedings of IEEE Workshop on Machine Learning for Signal Processing*, Sao Luis, Brasil, pp. 539-456, 2005.

[29] S. Zaferious, A. Tefas, I. Buciu, and I. Pitas, "Class-specific discriminant non-negative matrix factorization for frontal face verification," *Proceedings of 2005 International Conference on Advance in Pattern Recognition (ICAPR 2005)*, Bath, United Kingdom, 2005.