Journal of Computational Information Systems 5:1 (2009) 291-297
Available at http://www.Jofcis.com

IR and visible-light face recognition using canonical correlation analysis

Dianting LIU[1], Shungang HUA, Zongying OU, Jianxin ZHANG

China Key Laboratory for Precision and Non-traditional Machining Technology of Ministry of Education, Dalian University of Technology, China

Abstract

This paper proposes a novel multispectral feature extraction method according to the idea of canonical correlation analysis (CCA). Instead of extracting two groups of features with the same pattern (modality) as usual, the work explores another type of application of CCA that for extracting most correlated features from different face modalities to form effective discriminant vectors for recognition. Our goal is to search the complementary information in visible-light and infrared (IR) face imagery that are insensitive to the variation in expression and in illumination. Experimental results on Notre Dame face database show that the proposed CCA-based multispectral algorithm outperforms previous methods using visible-light imagery.

Keywords: Feature Extraction; Fusion; Canonical Correlation Analysis (CCA); IR; Visible Light; Face Recognition

1. Introduction

Face recognition, as a hot research area of biometric technologies during the past 30 years, has received considerable progress in controlled environments, where the lighting, pose, background, and quality of images were well controlled [1]. But in uncontrolled environments, random lighting sources from different directions significantly change visual appearances and influence the representation of the visible spectral images. This problem has been evidenced by the results of FRVT 2006, on which the best algorithm achieved a FRR (false reject rate) of 0.01 at a FAR (false accept rate) of 0.001 for matching facial images taken under controlled illumination, but in uncontrolled illumination experiments, the best FRR exceeded 0.10 at a FAR of 0.001 which was 10 times higher than the result got under controlled lighting condition [2]. To cope with the challenges posed by visible spectral images, thermal infrared images have been used as complements of visible light images for face recognition. Long-wave IR or thermal infrared images, captured in the range of 8–12 μm, represent the heat pattern of the object and are invariant to illumination and expression [3]. Compared with other spectral bands below the visible spectrum such as X-rays and ultraviolet radiation, thermal IR spectrum has no harm to the human body, and therefore can be employed in face recognition applications [4].

When it comes to problem about how to effectively fuse IR and visible features to improve recognition performance, the easiest way is to serially concatenate IR and visible feature vectors to form more informational vectors and has been proved very successful (see section 4). However, does any more effective method exist that could extract complementary information from IR and visible images and discard the useless? The main contribution of the paper is to explore the complementary information of multispectral face images adopting the idea of Canonical correlation analysis (CCA) to improve face recognition performance. CCA, which is proposed by H. Hotelling in 1936 [5], is one of the statistical methods finding basis vectors for two sets of variables such that the correlation between the projections of the variables onto these basis vectors are mutually maximized. It has the same importance as principal

[1] Corresponding author.
Email address: diantingliu@gmail.com (Dianting LIU)

January, 2009

component analysis (PCA) and linear discriminant analysis (LDA) in multivariate statistical analysis. In recent years, CCA has been applied to several fields such as signal processing, computer vision, neural network and speech recognition [6].

In the next section, the overview of the proposed method including the theoretical background of CCA used in feature extraction is presented. Section 3 and 4 indicates the description of experimental data and other experimental algorithms. Section 5 shows experimental results and necessary analysis. In Section 6, we make some concluding remarks.

2. Overview of the Proposed Method

Beginning with the idea that IR and visible imagery capture intrinsically different characteristics of the observed faces, we study the fusion of different groups of facial features which are most correlated. More formally, let us consider a training data set of K pairs of image vectors $(\boldsymbol{X},\boldsymbol{Y})=\{(\boldsymbol{x_1},\boldsymbol{y_1}),(\boldsymbol{x_2},\boldsymbol{y_2}),\ldots,(\boldsymbol{x_K},\boldsymbol{y_K})\}$ taking values in an n-dimensional image space and assume that each image pair belongs to one of L classes $\{C_1,C_2,\ldots,C_L\}$, where visible-light imagery corresponds to $\boldsymbol{X}$ in $(\boldsymbol{X},\boldsymbol{Y})$ and IR imagery corresponds to $\boldsymbol{Y}$ in $(\boldsymbol{X},\boldsymbol{Y})$. In order to address the problem of singular total scatter matrixes and take into account the recognition (discrimination) aspect of the data, Fisherfaces method [11],which is also called PCA+LDA, is employed prior to CCA.

In the paper, we proposed the following two steps for the construction of projection spaces of IR and visible face recognition:

Step 1 Transform the training set $(\boldsymbol{X},\boldsymbol{Y})$ respectively, into features spaces of lower dimensionality, as $\boldsymbol{X}\rightarrow\widetilde{\boldsymbol{X}}=\boldsymbol{P_X}^{\boldsymbol{T}}\boldsymbol{X}$ and $\boldsymbol{Y}\rightarrow\widetilde{\boldsymbol{Y}}=\boldsymbol{P_Y}^{\boldsymbol{T}}\boldsymbol{Y}$, where $\widetilde{\boldsymbol{X}}$ and $\widetilde{\boldsymbol{Y}}$ are dimension reductions using Fisherfaces for $\boldsymbol{X}$ and $\boldsymbol{Y}$ respectively, $\boldsymbol{P}_X$ and $\boldsymbol{P}_Y$ are projection matrices.

Step 2 Extract the maximum canonical correlation $\hat{\boldsymbol{X}}=\boldsymbol{\alpha}^{\boldsymbol{T}}\widetilde{\boldsymbol{X}}$ and $\hat{\boldsymbol{Y}}=\boldsymbol{\beta}^{\boldsymbol{T}}\widetilde{\boldsymbol{Y}}$ from $\widetilde{\boldsymbol{X}}$ and $\widetilde{\boldsymbol{Y}}$ based on the idea of CCA, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are two linear projection matrices.

In the first step, two projection matrices $\boldsymbol{P}_X$ and $\boldsymbol{P}_Y$ are learned from the training sets $(\boldsymbol{X},\boldsymbol{Y})$, to reduce the dimensionality from n to $L-1$ by Fisherfaces, which is achieved by using PCA to reduce the dimension of the feature space to $n-L$ firstly, and then applying the standard LDA to reduce the dimension to $L-1$.

In the second step, CCA is used to find two sets of basis vectors $\boldsymbol{\alpha}=[\boldsymbol{\alpha_1}\ \boldsymbol{\alpha_2}\ \ldots\boldsymbol{\alpha_d}]$ and $\boldsymbol{\beta}=[\boldsymbol{\beta_1}\ \boldsymbol{\beta_2}\ \ldots\boldsymbol{\beta_d}]$, one for $\widetilde{\boldsymbol{X}}$ and the other for $\widetilde{\boldsymbol{Y}}$ (d is the dimension of CCA subspaces), such that the correlations between the canonical variates $\hat{\boldsymbol{x}}_{\boldsymbol{i}}=\boldsymbol{\alpha}^{\boldsymbol{T}}\widetilde{\boldsymbol{x}}_i$ and $\hat{\boldsymbol{y}}_{\boldsymbol{i}}=\boldsymbol{\beta}^{\boldsymbol{T}}\widetilde{\boldsymbol{y}}_{\boldsymbol{i}}$ onto these basis vectors are mutually maximized, where $i=1,2,\ldots,d$.

This is done by maximizing the following correlation

$$\boldsymbol{\rho}=\frac{\boldsymbol{E}[\boldsymbol{\alpha}^{\boldsymbol{T}}\widetilde{\boldsymbol{X}}\widetilde{\boldsymbol{Y}}^{\boldsymbol{T}}\boldsymbol{\beta}]}{\sqrt{\boldsymbol{E}[\boldsymbol{\alpha}^{\boldsymbol{T}}\widetilde{\boldsymbol{X}}\widetilde{\boldsymbol{X}}^{\boldsymbol{T}}\boldsymbol{\alpha}]\boldsymbol{E}[\boldsymbol{\beta}^{\boldsymbol{T}}\widetilde{\boldsymbol{Y}}\widetilde{\boldsymbol{Y}}^{\boldsymbol{T}}\boldsymbol{\beta}]}}=\frac{\boldsymbol{\alpha}^{\boldsymbol{T}}\boldsymbol{E}[\widetilde{\boldsymbol{X}}\widetilde{\boldsymbol{Y}}^{\boldsymbol{T}}]\boldsymbol{\beta}}{\sqrt{\boldsymbol{\alpha}^{\boldsymbol{T}}\boldsymbol{E}[\widetilde{\boldsymbol{X}}\widetilde{\boldsymbol{X}}^{\boldsymbol{T}}]\boldsymbol{\alpha}\boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{E}[\widetilde{\boldsymbol{Y}}\widetilde{\boldsymbol{Y}}^{\boldsymbol{T}}]\boldsymbol{\beta}}} \tag{1}$$

Now observe that the covariance matrix of $(\widetilde{\boldsymbol{X}},\widetilde{\boldsymbol{Y}})$ is

$$\boldsymbol{C}(\widetilde{\boldsymbol{X}},\widetilde{\boldsymbol{Y}})=E\left[\begin{pmatrix}\widetilde{\boldsymbol{X}}\\ \widetilde{\boldsymbol{Y}}\end{pmatrix}\begin{pmatrix}\widetilde{\boldsymbol{X}}\\ \widetilde{\boldsymbol{Y}}\end{pmatrix}^T\right]=\begin{bmatrix}\boldsymbol{C}_{\widetilde{X}\widetilde{X}} & \boldsymbol{C}_{\widetilde{X}\widetilde{Y}}\\ \boldsymbol{C}_{\widetilde{Y}\widetilde{X}} & \boldsymbol{C}_{\widetilde{Y}\widetilde{Y}}\end{bmatrix}=\boldsymbol{C} \tag{2}$$

The total covariance matrix $\boldsymbol{C}$ is a block matrix where the within-sets covariance matrices are $\boldsymbol{C}_{\widetilde{X}\widetilde{X}}$ and $\boldsymbol{C}_{\widetilde{Y}\widetilde{Y}}$ and the between-sets covariance matrices are $\boldsymbol{C}_{\widetilde{X}\widetilde{Y}}=\boldsymbol{C}_{\widetilde{Y}\widetilde{X}}{}^T$

Hence, we can rewrite the function (1) as

$$\boldsymbol{\rho}=\frac{\boldsymbol{\alpha}^{\boldsymbol{T}}\boldsymbol{C}_{\widetilde{X}\widetilde{Y}}\boldsymbol{\beta}}{\sqrt{\boldsymbol{\alpha}^{\boldsymbol{T}}\boldsymbol{C}_{\widetilde{X}\widetilde{X}}\boldsymbol{\alpha}\boldsymbol{\beta}^{\boldsymbol{T}}\boldsymbol{C}_{\widetilde{Y}\widetilde{Y}}\boldsymbol{\beta}}} \tag{3}$$

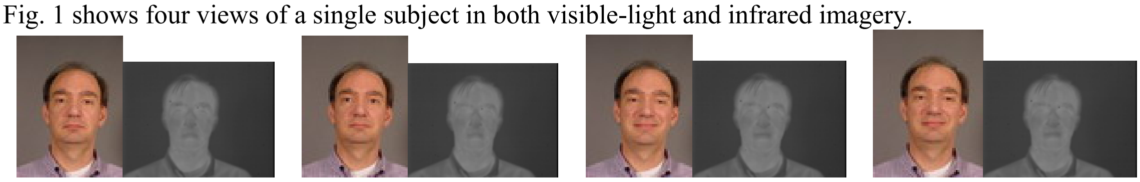
The maximum canonical correlation is the maximum of $\boldsymbol{\rho}$ with respect to $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. The subsequent canonical correlations are uncorrelated for different solutions, i.e.

$$\begin{cases} E[\boldsymbol{\alpha}_i^{\boldsymbol{T}} \widetilde{\boldsymbol{X}}\widetilde{\boldsymbol{X}}^{\boldsymbol{T}} \boldsymbol{\alpha}_j] = \boldsymbol{\alpha}_i^{\boldsymbol{T}} \boldsymbol{C}_{\widetilde{X}\widetilde{X}} \boldsymbol{\alpha}_j = 0 \\ E[\boldsymbol{\beta}_i^{\boldsymbol{T}} \widetilde{\boldsymbol{Y}}\widetilde{\boldsymbol{Y}}^{\boldsymbol{T}} \boldsymbol{\beta}_j] = \boldsymbol{\beta}_i^{\boldsymbol{T}} \boldsymbol{C}_{\widetilde{Y}\widetilde{Y}} \boldsymbol{\beta}_j = 0 \quad \text{for} \quad i \neq j \\ E[\boldsymbol{\alpha}_i^{\boldsymbol{T}} \widetilde{\boldsymbol{X}}\widetilde{\boldsymbol{Y}}^{\boldsymbol{T}} \boldsymbol{\beta}_j] = \boldsymbol{\alpha}_i^{\boldsymbol{T}} \boldsymbol{C}_{\widetilde{X}\widetilde{Y}} \boldsymbol{\beta}_j = 0 \end{cases} \tag{4}$$

As for recognition, first we project gallery image vectors $(\boldsymbol{x}_g, \boldsymbol{y}_g)$ and probe image vectors $(\boldsymbol{x}_p, \boldsymbol{y}_p)$ onto subspaces $\boldsymbol{P}_X$ and $\boldsymbol{P}_Y$ learned in step 1 as $(\widetilde{\boldsymbol{x}}_g, \widetilde{y}_g)$ and $(\widetilde{\boldsymbol{x}}_p, \widetilde{\boldsymbol{y}}_p)$ respectively. Then they are projected into the CCA subspaces as $(\hat{\boldsymbol{x}}_g, \hat{\boldsymbol{y}}_g)$ and $(\hat{\boldsymbol{x}}_p, \hat{\boldsymbol{y}}_p)$ using $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ found in step 2. The fused vectors $\boldsymbol{z}_g = \hat{\boldsymbol{x}}_g + \hat{\boldsymbol{y}}_g$ and $\boldsymbol{z}_p = \hat{\boldsymbol{x}}_p + \hat{\boldsymbol{y}}_p$ is used to calculate the correlation $\boldsymbol{s}$ as the matching score, where $\boldsymbol{s} = \boldsymbol{z}_g \cdot \boldsymbol{z}_p / (\|z_g\| \|z_p\|)$.

3. Description of Data Set

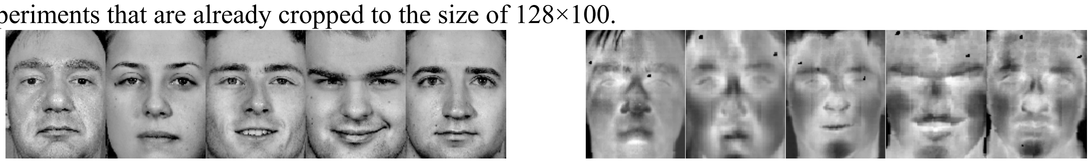
We presented and assessed the advantage of the proposed CCA-based multispectral face recognition scheme, using a data set from Notre Dame face database [12]. Specially, we used 1968 pairs of IR and visible face images corresponding to 63 subjects with variations in expression, lighting, and time lapse. Acquisitions were held weekly and most subjects participated multiple times across a number of different weeks. In the studio, three lights were located from left to right. One lighting configuration had the central light turned off and the others on. This is referred to as “LF’’. The other configuration has all three lights on; this is called “LM”. For each subject and illumination condition, two images were taken: one is with neutral expression, which is called “FA”, and the other image is with a smiling expression, which is called “FB”. Due to IR’s opaqueness to glass, all subjects were required to remove eyeglasses during acquisition. Fig. 1 shows four views of a single subject in both visible-light and infrared imagery.



(a) FA|LF (b) FA|LM (c) FB|LF (d) FB|LM

Fig. 1 Four views with different lighting and expressions in visible-light and IR imagery from Notre Dame face database

For the following experiments, we divided the IR and visible-light imagery into four groups named FA|LF, FA|LM, FB|LF and FB|LM respectively, and each group contains 492 pairs of IR and visible-light images from 63 subjects. Necessary image preprocessing has been done before the experiments. First, the centers of the eyes of an image are manually detected. Then rotation and scaling transformations align the centers of the eyes to predefined locations. Finally, the face image is cropped to the size of 128×100 to extract the facial region and histogram equalized. Fig. 2 shows some example images used in our experiments that are already cropped to the size of 128×100.



(a) (b)

Fig. 2. Examples images used in our experiments from Notre Dame face database: (a) samples of visible images and (b) samples of IR images under similar lighting and expression conditions as (a)

4. Description of Experimental Algorithms

For comparison purpose, four algorithms were evaluated:

Ⅰ FF(VI, IR)→CCA

Ⅱ FF(VI)+FF(IR)

Ⅲ FF(VI, VI)→CCA [13]

Ⅳ FF(VI) [11]

Algorithm Ⅰ is the proposed method described in section 2. VI is the shortened form of visible image and FF is the shortened form of Fisherfaces method.

Algorithm Ⅱ is to serially concatenate IR and visible feature vectors to form multispectral feature vectors. The training set $(\boldsymbol{X},\boldsymbol{Y})$ were firstly transformed to $(\widetilde{\boldsymbol{X}},\widetilde{\boldsymbol{Y}})$ by Fisherfaces as $\widetilde{X} = {P_X}^T X$ and $\widetilde{Y} = {P_Y}^T Y$. The projection of gallery sample $(\boldsymbol{x}_g, \boldsymbol{y}_g)$ and probe sample $(\boldsymbol{x}_p, \boldsymbol{y}_p)$ onto the projection matrices $\boldsymbol{P_X}$ and $\boldsymbol{P_Y}$ are $(\widetilde{\boldsymbol{x}}_g, \widetilde{\boldsymbol{y}}_g)$ and $(\widetilde{\boldsymbol{x}}_p, \widetilde{\boldsymbol{y}}_p)$ respectively. The projection reduces the dimensionality of $\widetilde{\boldsymbol{x}}_g, \widetilde{\boldsymbol{y}}_g, \widetilde{\boldsymbol{x}}_p, \widetilde{\boldsymbol{y}}_p$ from 12800 to 62. The serial combined feature vectors are defined by

$$\boldsymbol{z}_g = \begin{bmatrix} \widetilde{\boldsymbol{x}}_g \\ \widetilde{\boldsymbol{y}}_g \end{bmatrix} \tag{5}$$

$$\boldsymbol{z}_p = \begin{bmatrix} \widetilde{\boldsymbol{x}}_p \\ \widetilde{\boldsymbol{y}}_p \end{bmatrix} \tag{6}$$

Finally, the correlation of serial combined feature vectors $\boldsymbol{z}_g$ and $\boldsymbol{z}_p$ is adopted as the matching score.

Algorithm Ⅲ employs original visible imagery and its low resolution version as input of CCA. The object of introducing Algorithm Ⅲ here is to compare the classific information contained in low resolution imagery and IR imagery in the context of CCA. We first implement Haar wavelet transform onto the original visible images to obtain lower-resolution (32×25) version $\boldsymbol{X} = \{\boldsymbol{x} \mid \boldsymbol{x} \in \Re^{800}\}$, which corresponds to $\boldsymbol{X}$ in $(\boldsymbol{X},\boldsymbol{Y})$ in section 2 (low-resolution image is known as containing shape information of its high-resolution version); the original visible imagery corresponds to $\boldsymbol{Y}$ in $(\boldsymbol{X},\boldsymbol{Y})$. Then two steps are performed in the aforementioned formulations in section 2 to construct the projection spaces.

Algorithm Ⅳ is the standard Fisherfaces method using the original visible imagery $\boldsymbol{X}$ of the training set. The following projection reduces the dimensionality of gallery samples and probe samples from 12800 to 62.

5. Experimental Results and Analysis

In order to test the advantage of the multispectral CCA-based method in expression or illumination variation, two separate experiments are conducted. In the first experiment, the gallery and probe images own same expression but different illumination conditions; whereas in the second experiment the gallery and probe images share same illumination but different expressions.

In our experiments, recognition performance is characterized by two statistics: verification and false accept rates. The false accept rate is computed from comparisons between faces of different people. These comparisons are called non-matches. The verification rate is computed from comparisons between two facial images of the same person. These comparisons are called match scores. In addition, equal error rate (EER) is computed for the performance evaluation, too. EER is the error rate where the FAR (false non-match rate) and FRR (False match rate) assume the same value.

5.1. Experiments with Variations in Illumination

In this series of experiments, four data subsets corresponding to four sub-experiments were constructed from the database, each comprising one training set, one gallery and one probe. In each sub-experiment, the gallery and probe share variable lighting conditions but the same expression. Table 1 shows the experimental assignment in details.

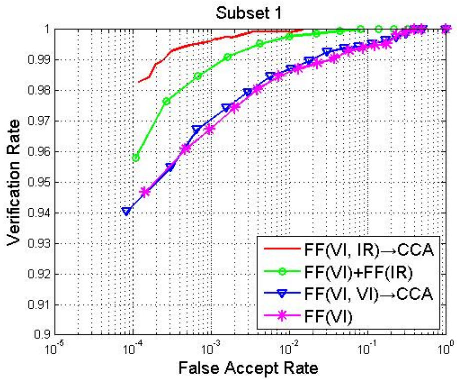
Table 1 Scheme of 4 experimental data subsets with variations in illumination

Subset	1	2	3	4
Training	FA|LF	FA|LM	FB|LF	FB|LM
Gallery	FB|LM	FB|LM	FA|LF	FA|LF
Probe	FB|LF	FB|LF	FA|LM	FA|LM

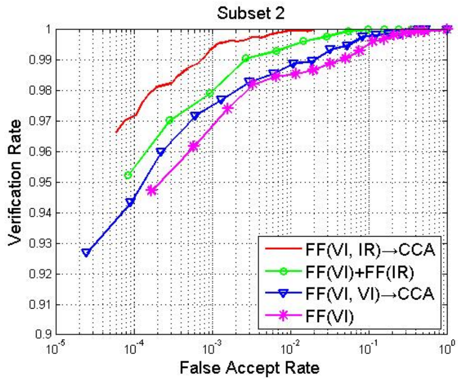
Table 2 provides EER of the four algorithms on 4 sub-experiments (in percentage) and Fig. 3 shows the receiver operating characteristic (ROC) for the results of four algorithms.

Table 2 Comparison of EER for the four algorithms on 4 sub-experiments in illumination variation (in %)

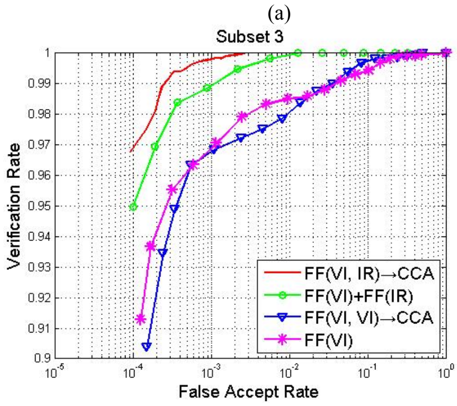
Equal Error Rate (%)		Subset				
		1	2	3	4	Average
Algorithm	Ⅰ	0.2135	0.2889	0.1396	0.1275	0.1924
	Ⅱ	0.4327	0.6606	0.2818	0.3728	0.4370
	Ⅲ	1.2050	1.1023	1.5810	1.2255	1.2784
	Ⅳ	1.2991	1.3964	1.4281	1.4347	1.3896



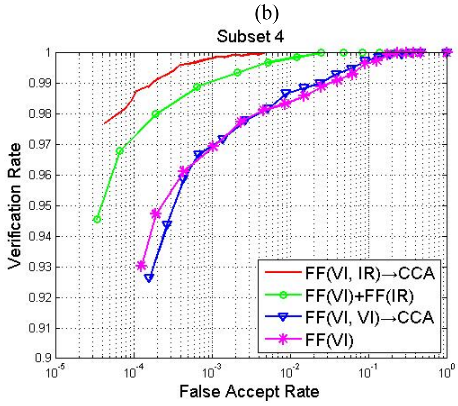
(a)



(b)



(c)



(d)

Fig.3 ROC curves for the four algorithms on 4 sub-experiments

In all of 4 sub-experiments, Algorithm Ⅰ gives the highest verification rate and the lowest EER of all; Algorithm Ⅱ performs next to Algorithm Ⅰ; Algorithm Ⅲ and Ⅳ achieve similarly worse performances. We can learn from Table 2 and Fig. 3 that in different lighting conditions, the proposed CCA-based algorithm (Algorithm Ⅰ) extracts IR and visible complementary information for classification and outperforms the method of serially combining IR and visible features (Algorithm Ⅱ). In the comparison of performance of Algorithm Ⅰ and Ⅲ, we find that IR imagery could provide more classific information than the low resolution visible imagery with respect to the CCA-based algorithm. We can also notice that low resolution visible imagery in Algorithm Ⅲ offers limited improvement on the experimental results when comparing with the performance of Algorithm Ⅳ.

5.2 Experiments with Changes in Expression

In order to evaluate the four algorithms' performance where faces are subject to expression variations, we adopted a similar series of sub-experiments as section 5.1. Four data subsets corresponding to four sub-experiments were constructed from the database. In each sub-experiment, the gallery and probe share different expression conditions but the same illumination. Table 3 shows the experimental assignment in details.

Table 3 Scheme of 4 experimental data subsets with changes in expression

Subset	1	2	3	4
Training	FA|LF	FB|LF	FA|LM	FB|LM
Gallery	FA|LM	FA|LM	FA|LF	FA|LF
Probe	FB|LM	FB|LM	FB|LF	FB|LF

Table 4 provides EER of four algorithms (in percentage) on 4 sub-experiments and Fig. 4 shows the ROC curves for the four algorithms when gallery and probe have different expressions.

Table 4 Comparison of EER for the four algorithms on 4 sub-experiments in expression variation (in %)

Equal Error Rate (%)		Subset				
		1	2	3	4	Average
Algorithm	Ⅰ	0.0777	0.2455	0.3544	0.1527	0.2076
	Ⅱ	0.3390	0.7874	0.6342	0.3187	0.5198
	Ⅲ	0.7173	1.3863	1.1194	1.1335	1.0891
	Ⅳ	0.9350	1.6407	1.9635	1.4978	1.5093

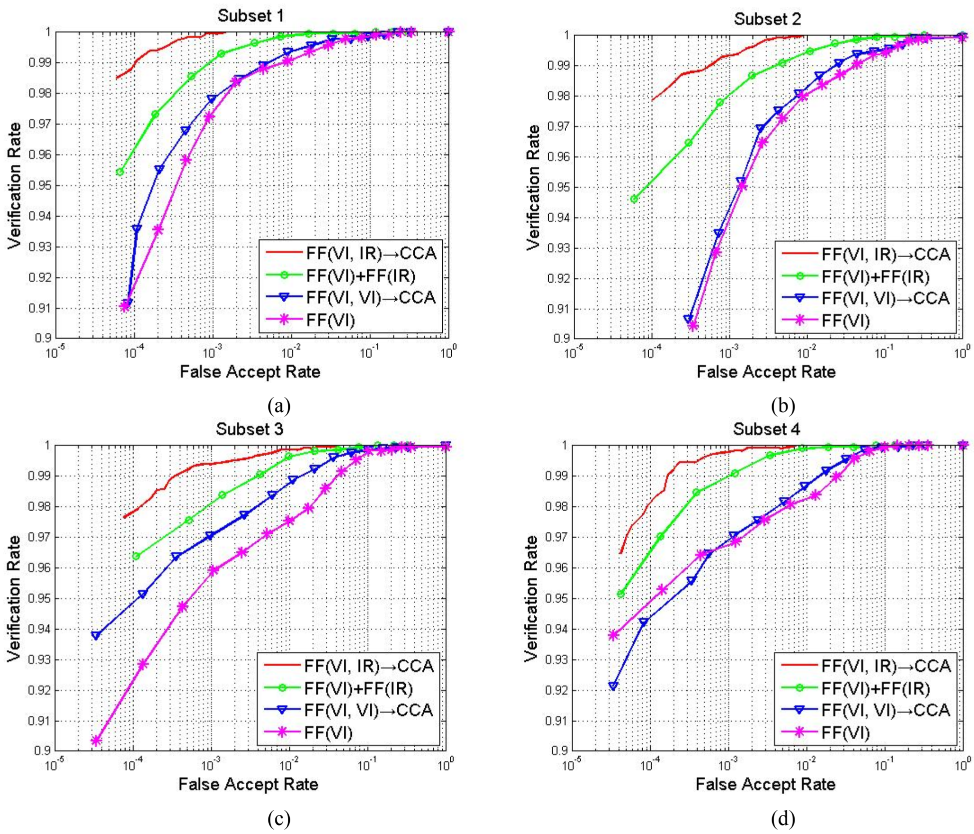


Fig.4 ROC curves for the four algorithms on 4 experimental subsets

In Table 4, Algorithm Ⅰ gains the lowest EER of all, and its average EER is about 60% lower than algorithm Ⅱ. From Fig. 4, we find that Algorithm Ⅰ also obtains the obvious advantage in verification rate at FAR=0.001 and FAR=0.0001 which reveals that Algorithm Ⅰ indeed extract effectively complementary information of IR and visible images that suitable for classification in variable expression condition.

6. Conclusions

In this work, we presented a new CCA-based multispectral algorithm for extracting and fusing IR and visible features for the purpose of face recognition. Our objective is to utilize the complementary information in the two spectra and improve recognition performance across variable lighting and expressions. The experimental results obtained on Notre Dame face database are encouraging: the lowest EER and the highest verification rate were gained by the proposed CCA-based method in both variable illumination and expression conditions. It is observed that employing CCA-based feature extraction algorithm is quite effective in combating the detrimental effects of illumination and expression variations.

Acknowledgement

The authors would like to thank Professor Patrick Flynn and Dr. Xin Chen for providing University of Notre Dame biometrics multimodal databases. This work is supported by Photoelectron+X interdisciplinary research funds of Dalian University of Technology (2006), and research funds of Dalian University of Technology and Shenyang Institute of Automation Chinese Academy of Sciences joint research program.

Reference

[1] Zhao W, Chellappa R, Phillips P, and Rosenfeld A. Face recognition: a literature survey. ACM Computing Surveys, Vol. 35, No. 4, pp. 399–458, 2003.

[2] Phillips PJ, Scruggs WT, O'Toole AJ, *et al*. FRVT 2006 and ICE 2006 large-scale results. Technical report, National Institute of Standards and Technology, NISTIR 7408, March 2007. Available from: <http://iris.nist.gov/ice>.

[3] Singh R, Vatsa M, and Noore A. Integrated multilevel image fusion and match score fusion of visible and infrared face images for robust face recognition. Pattern Recognition, Vol. 41, pp. 880 – 893, 2008.

[4] Kong SG, Heo J, Boughorbel F, *et al*. Multiscale fusion of visible and thermal IR images for illumination-invariant face recognition. International Journal of Computer Vision, Vol. 71, No. 2, pp. 215–233, 2007.

[5] Hotelling H, Relations between two sets of variates. Biometrika, Vol. 28, pp.312-377, 1936.

[6] Borga M, Learning multidimensional signal processing, Linkoping studies in science and technology, Dissertations, Vol. 531, Department of Electrical Engineering, Linkoping University, Linkoping, Sweden, 1998.

[7] Yu SJ, Direct blind channel equalization via the programmable canonical correlation analysis, Signal Process, Vol. 81, pp. 1715–1724, 2001.

[8] Fyfe C, Lai PL. Canonical correlation analysis neural networks. International Conference on Pattern Recognition, Vol. 2, Barcelona, pp. 977–980, 2000.

[9] Choi HC, King RW. Speaker adaptation through spectral transformation for HMM based speech recognition. IEEE Int. Symp on Speech, Image Process and Neural Networks, Hong Kong, Vol. 2, pp. 686–689, 1994.

[10] Weenink D, Canonical Correlation Analysis, Institute of Phonetic Sciences, University of Amsterdam, Proceedings, Vol. 25, pp. 81–99, 2003.

[11] Belhumeur PN, Hespanha JP, and Kriegman DJ. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. IEEE Trans. Pattern Anal. Machine Intell., Vol. 19, pp. 711–720, Jul. 1997.

[12] Chen X, Flynn P, and Bowyer K. IR and visible light face recognition. Comput. Vision Image Understanding, Vol. 99, No. 3, pp. 332–358, 2005.

[13] Sun Quan-Sen, Zeng Sheng-Gen, Heng Pheng-Ann, Xia De-Shen. The Theory of Canonical Correlation Analysis and Its Application to Feature Fusion. Chinese Journal of Computers, Vol.28, No.9, pp.1524-1533, 2005.