A kind of Face Recognition Method Based on CCA Feature Information Fusion

NIU Li-Xia [1] and LI Guo [2]
1. Department of Business Administration, Liaoning Technology University, Huludao 125105, China
2. Department of Safety Inspection, Fuxin Xinda Security Engineering CO.LTD, Fuxin 123000, China

***Abstract*—In order to achieve more local facial feature, a kind of sub image face recognition method based on RS-Sp CCA feature information fusion is proposed in this paper. According to take samples for the local facial feature of sub image and use CCA to fuse the global facial feature and the local facial feature information after sampling, the global feature of image can be fully used to construct much more different kinds of component classifiers. Then, make experimental analysis on databases of 3 standard facial data sets. At last, the results show that sub image face recognition method based on RS-Sp CCA feature information fusion is better than the simple feature sampling method and feature fusion method. Besides, it can efficiently improve the face recognition rate.**

***Index Terms*—Feature Fusion, CCA, Facial Recognition, Classifier**

I. INTRODUCTION

As a focal point of research in the field of machine vision and pattern recognition, the face recognition gets more and more attention in recent decades. In face recognition, face feature description is one of the key steps [1-3]. In general, face features are divided into global features and local features. In these features, global features mainly describe facial overall properties (such as color) , which is used for rough matching, while local features mainly describe facial details change (such as scar) , which used for detailed confirmation. A large number of global and local feature extraction methods have been proposed, including PCA, LDA, LPPP and the other classical global feature methods. PCA obtains the features with the optimal reconfiguration performance according to maximizing sample dispersion; LDA obtains the features with the optimal discriminant ability according to the radio between the minimizing intra-class scatter and inter-class scatter; LPP obtains the features with the potential discriminant ability according to the most largely reserve the geometry among the samples. However, when they are getting global features, they lack of insensitivity and robustness to local changes (such as shade and light). In order to get more local features and try to overcome the sensitivity to the local changes, the methods based on local features are widely used in face recognition, such as Gabor wavelet method, LBPP and sub image method, etc [4-8]. Gabor wavelet method magnifies eyes, nose and mouth and some other local features by extracting multi-scale and multi-direction spatial frequency domain features within the specific area of image [9]. LBP obtains the local texture information according to make binary coding for local region, while the sub image method is to make division for the original image, and then separately extract features in each of the sub image to obtain local features.

Because the sub image method can effectively solve the problem of small samples, which improves the robustness of algorithm for light, shade and the other change factors, besides it is simple, intuitive and easy to popularize to other global or local feature extraction method, in recent years, the method based on sub image gets a lot of attention [10-13].

According to extract and use local information of face image, Gottumukkal proposes module PCA (mPCA) method. In the mPCA, firstly, face is divided into several blocks of equal size, and then sub image set which is formed by all child blocks is looked on as new training set. Finally, execute PCA in this new training set to extract features. Almost at the same time, Chen and Zhu proposes model sub-module PCA (SpPCA) [14-16]. Unlike mPCA, SpPCA respectively extracts local features from each sub image, and then all acquired local features are gotten together to form the final global features. Later then, Tan and the others propose the adaptive weighted Aw - SpPCA method. The method constructs a classifier on each sub image set, and then respectively makes classification for corresponding sub-blocks of unknown samples. Finally, use weighted voting method to fuse all classification results. On the basis of the literature, a lot of work about the sub image has been published. These methods have something in common that they all obtain sub image features through a particular method, and then construct a classifier on each sub image and they all acquires some degree of success. However, the classifiers that they construct are all simply based on local features or global features, and don’t make full use of the relationship among local features or between global and local features. In face recognition, the content and function that are described by global features and local features are different. The method based on global features tend to lack of robustness for illumination, facial expression and occlusion of and so on, while the simple method based on local features ignores the relationship between the local features so that it loses a part of global features. Therefore, the fusion of global and local features has become one of research directions in face recognition.

doi:10.4304/jmm.8.5.503-511

Recently, Hong and the others propose a kind of face recognition method SpCCA based on feature fusion. This method divides the face image into several blocks, and separately extracts the global features and local features from the whole face image and each sub image, and then use canonical correlation analysis (CCA) to fuse the global features and local features. Finally, voting method is used for final decision. Compared with traditional sub image method, SpCCA classifier fuses the relevant local and global feature, which makes more full use of the complementary relationship between the global features and local features. All in all, the method obtains the better performance.

Sub image method constructs a component classifier for each sub image, and then uses the integration of multiple classifiers to form the final decision, so, essentially it is multiple classifier systems. However, the number of the component classifier and diversity between each classifier is completely decided by the sub image classification form. Therefore, when the classification form of sub image changes, its performance may be severely affected. Random subspace method (RMS) is a kind of important method building variety component classifiers. It doesn't suffer from any constraints and it randomly selects features from the entire features so as to obtain different feature subset and train different classifiers on the subset. At present, the stochastic subspace method is widely used in face recognition. Wang and Tang firstly apply the stochastic subspace method to face recognition field; Zhang and the others construct optimal stochastic subspace dimension for discriminant analysis method. We apply the feature random sampling to face sub image and Semi-RS method is proposed. Semi - RS not only wins the more superior performance than the literature [6], but also it effectively increases the diversity between the component classifiers.

Inspired by SpPCA and Aw-SpPCA, this paper proposes a kind of face recognition method based on CCA feature information fusion (RS - SpCCA) and applies it on face recognition. This method fuses the local and global information through canonical correlation analysis (CCA). The first step is to take the correlation between global and local features vector as effective discriminant information, which not only reaches the purpose of the fusion of local and global information, but also eliminates the information redundancy between the features. Based on the lower price, at the same time, it realizes the advantages of the above two kinds of feature extraction. The second step is to construct more diverse component classifiers through the feature sampling and make full use of the relationship between global features and local features according to feature fusion. In the AR, Yale, ORL databases, the experimental results show that the RS-SpCCA method based on the feature sampling and information fusion is better than simple information fusion method and feature sampling methods (Semi - RS). What is the most important is that compared with SpCCA method, the sensitivity of RS-SpCCA for the size of sub image is smaller and it generalizes the scope of application of the relevant recognition analysis of typical feature information fusion.

This paper mainly does some developing and innovative work in the following aspects:

(1) Analyze the method of image feature fusion. To make a detailed analysis for several steps of RS - SpCCA method, in order to get more local facial characteristics, through sampling their local characteristics of subimage, classify the unknown information. And then using CCA global and the image characteristic information fusion after sampling, it can make full use of the global characteristic of image, so as to build up more different kinds of component classifier. This method effectively fuses local and global information characteristic, it compares with traditional method, not only reduce cost, but also effective discriminant information, and eliminate information redundancy between the characteristics, therefore, to realize the effective recognition of the unknown images.

(2) Carry on the experimental result testing analysis for the AR, Yale, ORL three data sets. Among them, the AR data sets are used to test the method's performance in the case of time change, shelter etc, Yale for testing the stability of RS-SpCCA under different training sample set, ORL is used to test the recognition in case of image slightly posture change. The experimental results show that this method's performance is better than simple sampling method and characteristics fusion method, face recognition rate is improved effectively.

(3) Discuss and analyze RS - SpCCA method. Randomly sampling each image in the collection and build more component classifier and parameter selection. In addition, to make a recognition effect performance test for collected images and for its result make contractive analysis. Finally, for SpCCA and Semi - RS method, the result shows that RS - SpCCA in training and testing phase with high time complexity. And the speed of calculation identification is faster and the effect is better.

II. IMAGE FEATURE FUSION METHOD

On the basis of sub image, RS-SpCCA introduces characteristics sampling and the ideal of information fusion at the same time, it wants to build more Component Classifier and make full use of the relationship between local features and global features. RS - SpCCA methods can be divided into the following steps:

(1) Face image is divided into sub modules, and regard it as the local characteristics vector of CCA fusion;

(2) To carry on random sampling of characteristics for each sub image;

(3) To extract the overall features of the original face image, and regard it as global feature vector badge of CCA fusion;

(4) Use the feature of sampling from sub image and CCA global feature, to carry on information fusion

(5) Classify unknown images, to obtain recognition results.

The detail description of this method as below .

A. The Division Method of Sub Image

Usually, there are two different kinds of classification methods: one kind is partial organ method, another method is local area method. Local organs carry on division mainly according to the feature organ of human face, such as eyes, mouth, nose area and so on. Owing to this division often relies on a certain characteristic test method or some degree of human intervention; therefore, it has certain restrictions on operation; The local area is according to the specific coordinates of image to divide, because this method often can get better recognition result than local organs method, so this article use the most simple, don't overlap and equal size's rectangular area to carry on division of the image.

Set the original training set containing N image that size is m×n. First of all to divide each image according to the same way to split it into L sub image that the size is equal (the size of each sub image is p = m by N/L) ; Then all images in the same position of the sub image is converted into a column and combine, forming a sub image set. Since there are L sub images of every image, therefore, for the whole training set, it produced L sub image training sample collection $\{T_i\}_{i=1}^{l}$. The process of construction as shown in figure 1.

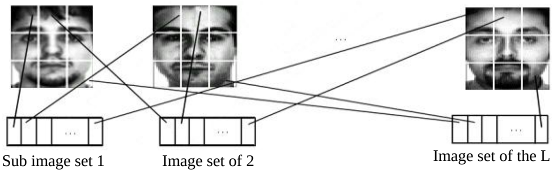


Figure 1. The construction of sub image

B. Random Sampling of Image Characteristics

RS - SpCCA performs random sampling on each image set, therefore, for the i_{th} sub image set Ti, based on the ideal of random subspace method, from 1 to p (p for the sub image size) randomly generate $p^*(p^* < p)$ index vector $v_i^1 = \{I_1^I, I_2^I, ..., I_N^I\}$, and then according to V_1^I, extract features from T_i and it expressed as T_i^1 this process repeated independently K times (K as sub image set classifier) , accordingly gained K feature subset $T_i^k(k = 1, 2, ..., k)$.

C. Global Feature Extraction

There are multiple ways of Global feature extraction, PCA, LDA and LPP all are classic global feature extraction method. In this paper, we adopt PCA global feature extraction. When actual apply this method, if you find more suitable overall feature extraction method or with some special needs, it only needs to replace CCA the second group characteristic vector.

For the training set T_r, it uses PCA feature extraction, the projection matrix under w_{PCA}, is available, the T_r samples uses w_{PCA}, projection, to obtain a group of low dimensional characteristics $T_{T_u} = (y_1, y_2, ..., y_n) m_2 \times n$. When extract the overall information, it is necessary to pay attention to the numbers of dimensions of the overall information extraction as well as local information should be restricted, to meet the dimensions $m_2 \le n - c$ of the global information, which ensures the overall covariance matrix composed by CCA in the second group of feature vector is a nonsingular. Therefore, overall covariance matrix that composed by two groups of characteristic vector of CCA all are nonsingular, so as to solve the small sample problem.

D. CCA Information Fusion

Receptively with CCA to fuse each train set pair (T_r^1, T_r^2) . According to local characteristics, can obtain L independent sub training set $T_r^1, T_r^2, ..., T_r^l$, for each training set $T_r^i (i = 1, 2, ..., L)$ N samples, C category, sample dimension. That is $Tr^i = (x_1^i, x_2^i, ..., x_n^i) m \times n$, and $T_{r_y} = (y_1, y_2, ..., y_n) m^2 \times n$.

Using training set pair (T_r^i, T_{r_y}) construct CCA projection matrix pair (w_x^i, w_y^i) , it asks through the projection matrix after projection on the original sample obtain the greatest correlation between two groups of low dimensional characteristics, and each component between two groups of low-dimensional features are not related. In this way, we can extract the nature characteristics of the same sample in different representation mode, while ignoring some redundant features because of different patterns, remove the same characteristics in the different representation model, so to improve the recognition rate but also improves the computational complexity.

Define the covariance matrix $s_{xx}^i = E[xx^T]$, of first group characteristic vector T_r^i , covariance matrix $S_{YY} = e[YY^t]$, of the second set of characteristic vector T_{T_Y} . Between them the mutual association matrix as $S_{XY}^I = e[XY^t]$, the optimal projection matrix is defined as:

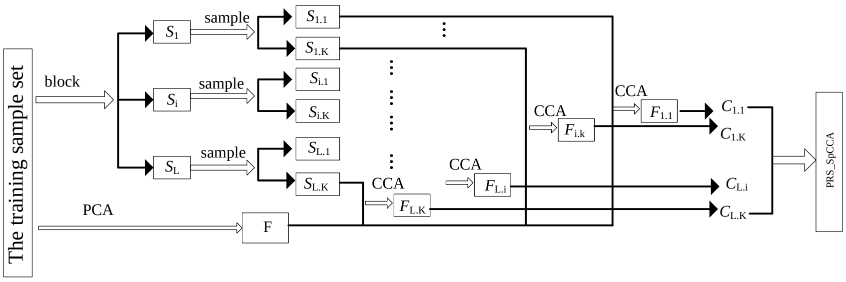
$$\left[w_x^i, w_y^i\right]_{opt} = \left[\left(w_{x1}^i, w_{x2}^i, ... w_{xr}^i\right), \left(w_{y1}^i, w_{y2}^i, ... w_{yr}^i\right)\right]$$
$$= \arg \max_{w_{xk}^i, w_{yk}^i} \frac{w_{xk}^{i\ T} s_{xy}^i w_{yk}^i}{\sqrt{w_{xk}^{i\ T} s_{xx}^i w_{xk}^i . w_{yk}^{i\ T} s_{yy} w_{yk}^i}}, 1 \le r \le c - 1 \tag{1}$$

Meet

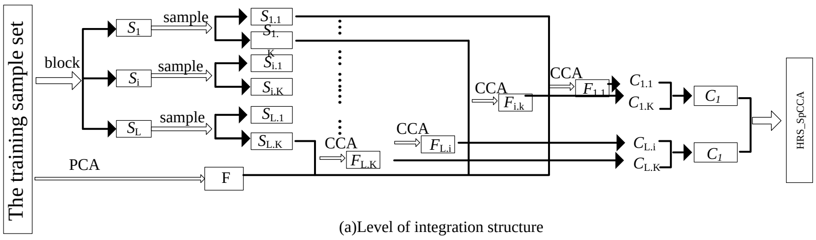
$$\begin{cases} w_{xk}^{i\ T} s_{xx}^i w_{xk}^i = w_{yk}^{i\ T} s_{yy}^i w_{yk}^i = 1, \\ w_{xk}^{i\ T} s_{xx}^i w_{xj}^i = w_{yk}^{i\ T} s_{yy}^i w_{yj}^i = 0, \\ j, k = 1, 2, ..., r, k \ne j \\ w_{xk}^i \in R^m, w_{yj}^i \in R^{M2} \end{cases} \tag{2}$$

According to the method of CCA, can known:

$$\begin{cases} s_{xy}^i s_{yy}^{-1} s_{yx}^i w_x^i = s_{xx}^i w_x^i \wedge i \\ s_{yx}^i (s_{xx}^i)^{-1} s_{xy}^i w_y^i = s_{yy}^i w_y^i \wedge i \end{cases} \tag{3}$$



(a)Parallel integration structure



(a)Level of integration structure

Figure 2. The structure diagram of RS-SpCCA

Among them, A is the diagonal matrix that is generalized eigenvalue before problem (5) the largest nonnegative eigenvalue in descending order, the column of 吡 and 叫 is the eigenvectors that the eigenvalue correspond to and meet type (4).

All sub training set $Tr^1, Tr^2, ..., Tr^L \ and \ Tr_y$, respectively as CCA projection matrix are constructed above. Then can get a group of projection matrix pair $(w_x^1, w_y^1), (w_x^2, w_y^2), ..., (w_x^L, w_y^L)$. that used to features extraction. Because of CCA's first group characteristic vectors is each local characteristics, the feature vector of second group used global characteristic, after CCA `s projection, each vector of the first group of features `s L vector turn the local characteristics of a piece is given priority to, blends the overall characteristics, each vector of second group of feature `s L vectors turn a global characteristics is given priority to, combined with the local characteristics of a piece. In this article use $t^T w_t$, namely, the local features as discriminant vector. In the same way, using CCA to fuse each element of $\{T_i^k\}_{i=,2,...l,k=1,2,...k}$ and F respectively, can obtained a set of discriminant vectors can be used for classification.

E. Classification of Unknown Image

When identifying unknown image Y, first should divide the image according to the step (1), and get L sub image $y_i = (I = 1, ...L)$. Then use the generated index vector V_I^K. in section 1.2 for Y_i, sampling , obtained characteristics y_i^k after the sampling, and then project y_i^k to w_i^k obtain the relevant features $(y_i^k)^T * w_i^k$. That is fused global features, last use nearest neighbor classifier to classify $(y_i^k)^T * w_i^k$. Due to each sub image need K times in sampling, therefore, classifying the sub image I, produce K classification results; and for the unknown image, there is L * K classification results. For different purposes, we respectively component classifier using different ways of combination:

1) From the perspective of diversity of component classifiers, each classifier adopts the way of parallel combination;

2) From the perspective of improving the performance of sub image classifier, each classifier adopts the way of hierarchical combination.

Two kinds of combination way as shown in figure 2

Now define symbol $d_c^{i,k} \in \{0,1\}, when \ d_c^{i,k} = 1$, represents the first K times sampling characteristic x_i^k of the ith sub image belongs to the class c; $d_c^{i,k} = 0$ does not belong to the class c. So for the parallel RS - SpCCA - SpCCA (PRS) structure, the classification result can be represented as:

$$identity(y) = \arg \max_{1 \le c \le c} \left(\sum_{i=1}^{l} \sum_{k=1}^{k} d_c^{i,k} \right) \qquad (4)$$

While for the level of the RS - SpCCA (hereinafter referred to as HRS - SpCCA) structure, actually, it performed twice integration. HRS - SpCCA firstly within each sub image for an integrated (voting) to determine the corresponding category, and then make the second combination for L sub image (voting) to get final decision. Twice integration (HRS – Sp CCA) results can be represented as:

$$identity(y) = \arg \max_{1 \le c \le c} \sum_{i=1}^{l} \left(\arg \max_{1 \le c \le c} \left(\sum_{k=1}^{k} d_c^{i,k} \right) \right) \qquad (5)$$

Because RS - SpCCA can effectively improve the accuracy of SpCCA component classifier or raise the diversity between the component classifier, therefore, in the intuition, RS - SpCCA can obtain better performance than SpCCA. In addition, in order to make more full use of the diversity among different classifiers, it would be a better choice to use parallel combination as shown in figure 2 (a). The experimental results also show that the PRS - SpCCA has a better performance than HRS - SpCCA. Similar results that parallel structure has better performance than the hierarchy has gotten validation in the literature [8].

RS-SpCCA makes random feature sampling for each sub image set, and then the sampling feature subsets are taken as the fusion of global features and local features. Obviously, when random sampling rate is 1, parallel RS-SpCCA and hierarchical RS-SpCCA all degrades to be SpCCA. Thus, SpCCA is just a special case of RS - SpCCA.

III. Analysis of Experimental Results

In order to evaluate the performance of RS-SpCCA, we make experiment on the three datasets, AR, Yale and ORL. In these datasets, the AR data sets are used to test the performance of this method in the case of time change, shelter and the others. Yale is uses to test the stability of RS-SpCCA under different training sample sets. ORL is used to test the recognition ability when the test image has slight gesture changes.

Because the literature [12] has made a comprehensive comparison among SpCCA, Eigenfaces, Fisherfaces, PCA + CCA, Aw - SpPCA and SpPCA ; the literature [7] has carried on the comparison for Semi-RS and M-Eigenfaces , Nitesh RS ' and so on, and the experiment proves that the SpCCA and Semi - RS are superior to the ratio method, in this paper, we only compare the RS-SpCCA with SpCCA and Semi - RS. In experiment, according to the experience conclusion (when sub image is around 1/100 of the original image size, the sub image can achieve better effect) that is given in the literature [13] (i.e., SpCCA) , we make initial settings for the size of the sub image of each data set (as table 1). For the two methods based on random sampling Semi-RS and RS-SpCCA, we set that random sampling rate is 0.5 and the number of classifier of each image set is set to 20. Then, experiment is independently executed 10 times and takes average as the final recognition results.

TABLE I. RS-SpCCA SpCCA and Semi - RS parameter Settings

data set	Image size	Random sampling rate
AR	6×6	0.6
Yale	5×5	0.6
ORL	5×5	0.6

A. The Experiments on AR Dataset

AR database is a very challenging database, which contains more than 3, 200 frontal face images of 126 people (76 male and 50 female). Each person has 26 different images. In these images, the first 13 images are taken in period 1, and the after 13 images is taken in the second period after two weeks. The data used in the laboratory is provided by Martinetz , which contains 2 600 images of 100 people (50 female and 50 male). The original face image size is 165 x 120. In order to calculate data conveniently, we change the image to 66 x 88 sizes. Figure 3 shows the 26 images of one people. As the experiment setting in literature [7], choose the first 7 images in the first period as the training and according to the change, the other images are divided into seven different subsets (namely AR77: Session2 1 ~ 7; AR73Exp: Session2 2 ~ 4; AR73Illu: 5 ~ 7 in Session2; AR73SungS1:8 ~ 10; in Session1 AR73Scarf1: Session1 in 11 ~ 13; AR73SungS2: Session2 in 8 ~ 10; AR73Scarf2: Session2 in 11 ~ 13) so as to test the performance of the various methods under different conditions.

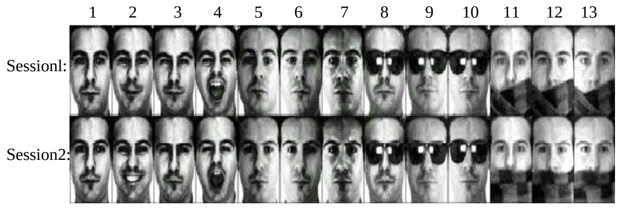


Figure 3. All samples of one people in AR database

1) The Changes of Light and Expression

The experiment uses AR73Exp and AR73Illu to evaluate the performance of RS-SpCCA method when light and expression change and compared with SpCCA and Semi-RS method, the experimental results are shown in table 2. From the table 2, we can see the follows.

1) PRS-SpCCA gains the best classification accuracy. Under the condition that recognition rate is very high, it still can improve at least one percent.

2) Because the PRS-SpCCA can make full use of the diversity between the classifiers, PRS-SpCCA is better than the HRS - SpCCA in overall and the similar conclusion has also embodied in Semi - RS method.

3) Either in hierarchical structure or in parallel structure, RS-SpCCA performs better than Semi-RS method having the same structure, which shows that fusing global feature is necessary.

TABLE II. Recognition precision (%) of the two test sets, AR73Exp and AR73Illu

	SpCCA	HSemi-RS	PSemi-RS	PRS-SpCCA	HRS_SpCCA
Exp73	96.65	96.60	95.78	97.89	97.89
Illu73	97.34	98.76	96.35	98.68	99.05

2) Image Block

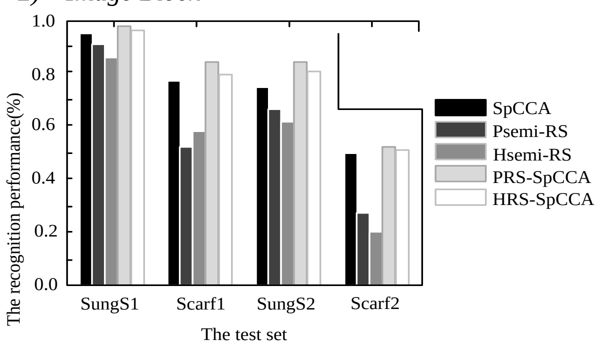


Figure 4. The recognition performance of RS-SpCCA on four test sets

Experiment uses AR73SungS1, AR73SungS2, AR73Scarf1 and AR73Scarf2 to test the RS - SpCCA recognition performance of images which are covered with dark glasses, scarf and the others. Figure 4 shows the experiment results of different methods. From the results, we can see that RS-SpCCA all gains the best performance in 4 test sets. Relative to the suboptimal method (not including HRS - SpCCA), the smallest increase amplitude is 3%, and the largest amplitude reaches 10%. These results strongly suggest that RS-SpCCA has higher robustness for serious barrier.

B. Experiment on the Yale Dataset

Yale database contains 165 pieces of gray images of 15 people and every people have 11 pieces of gray images. All images are frontal images and spatial location is relatively the same and there is almost no rotation change, but a few images have scale change and the light intensity has a little change. The images used in this experiment are from the data that ca_i processes (www.rjucadmg.cn/dengcai/Data/data.html). The size of images is 32x32. Experiment randomly selects the samples with a given number (5 or 6) from 11 samples of one of each type as the training and the other samples are taken as test. Experiment repeats 20 times and records the mean and variance.

The experiment results are shown in table 3. The data in the table shows that RS-SpCCA not only gets a better average precision, but also has the smaller variance. It suggests that the RS-SpCCA method has more stable performance than SpCCA and Semi-RS.

C. The Experiments on ORL data Set

ORL face data set contains 400 pieces of gray images from 40 individuals and each has 10 different pictures. Some of images in the data set are taken in different periods, and some images have changes in the facial expressions (open or close my eyes) , facial details (wear glasses or not) and the scale size (the largest scale of no more than 10%) , and some of images has the rotation which is not more than 20 degrees. The data used in the experiments is also handled by ca_i and the size of images is 32 x32. Experiment randomly selects 5 (or 6) samples from one of each type samples as the training and the other samples are taken as test. Experiment repeats 20 times and the average is taken as the final recognition accuracy. The experiment results are as shown in table 4. From the table, we can see both PRS-SpCCA and HRS – SpCCA show the better performance for mild attitude changes. Compared with the other methods, it at least increases one percent, but it's worth to express that when images have bigger posture change, because at this time, the extracted local features can't have too much effect on the correct recognition, the sub image methods include the recognition performance of Semi-RS and SpCCA, it may decrease along with the increase of gesture attitude change.

TABLE III. THE IDENTIFICATION PRECISION (%) IN ORL TEST SET

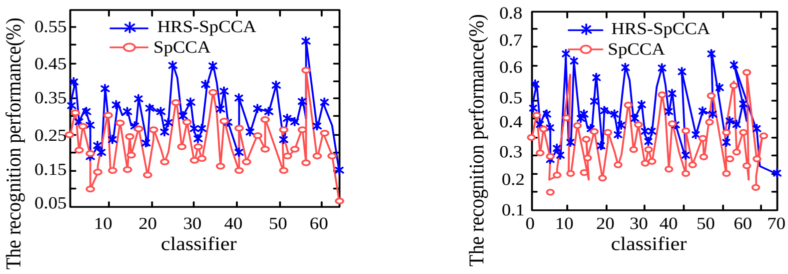
	SpCC A	HSemi -RS	PSemi-R S	PRS-SpCC A	HRS_SpCC A
G6	94.35	92.58	95.56	96.78	96.05
G5	92.84	93.37	94.23	96.13	94.35

D. Discussion and Analysis

From the perspective of each component classifier performance that improves sub image method (i.e., build on the sub image classifier performance), we use hierarchical structure to combine the L×K classifiers. That is to say, first of all, make combination for the K classifiers in each sub image which is as the classification results of sub image, and then make combination for the classification results of all sub images. In order to illustrate that constructing multiple classifiers in sub image can improve the classification ability of sub image, we respectively make comparative test on the ORL and Yale dataset (the first 6 samples of each kind are taken as training, and the others is taken as test) and make comparison between HRS - SpCCA and SpCCA. The results are shown in figure 5. From the figure, in addition to a few points, after the first vote, the classification ability of HRS - SpCCA sub image is obviously higher than that of SpCCA. Because the diversity between each classifier (between the sub image) basically remains the same, HRS-SpCCA can obtain better recognition performance than SpCCA. Another thing which is worth mentioning is that in the analysis of above experiment results, we don’t compare HRS-SpCCA and PSemi-RS. Its main reason is that they adopt two different classifier combination methods and there is no comparability. Therefore, when this paper shows the good performance of HRS-SpCCA, it doesn’t make comparison with PSemi - RS.

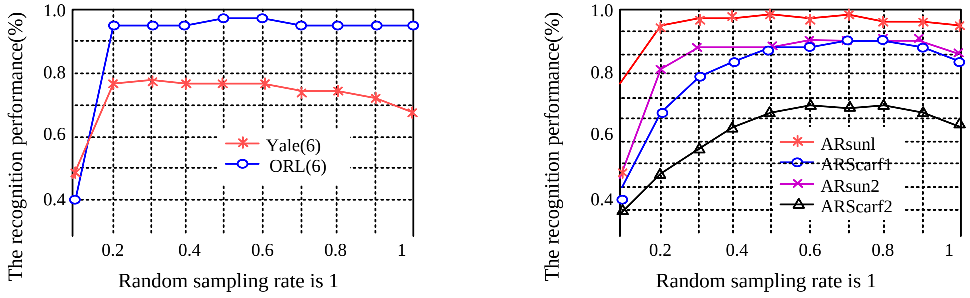
TABLE IV. THE AVERAGE PRECISION AND VARIANCE (%) IN THE YALE DATASET

	SpCCA	HSemi-RS	PSemi-RS	PRS-SpCCA	HRS_SpCCA
G6	66.25 ± 4.35	72.56 ± 3.25	75.56 ± 1.23	76.78 ± 2.83	74.05 ± 2.80
G5	60.67 ± 0.45	73.45 ± 1.8	69.59 ± 1.50	76.13 ± 0.56	68.35 ± 2.14



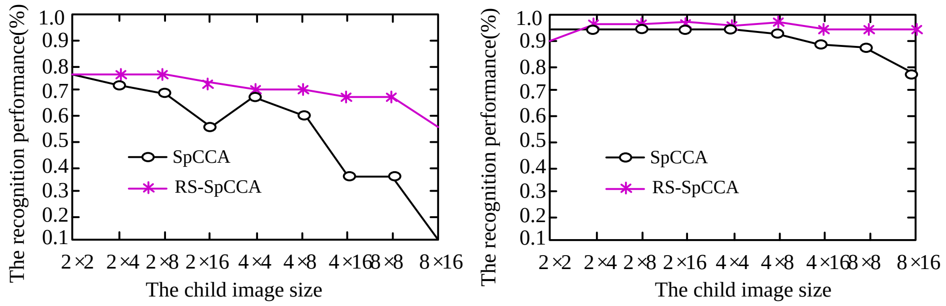
(a) Yale dataset (b) ORL data set

Figure 5. The comparison between each classifier in SpPCA and HRS – SpCCA



(a) ORL and Yale dataset (b) AR data set

Figure 6. The effect of random sampling rate on the performance of RS-SpCCA



(a) Yale dataset (b) ORL data set

Figure 7. The effect of different sub image sizes on the RS-SpCCA and SpCCA

E. Parameter Selection

RS-SpCCA performs random sampling in each sub image set, so compared with SpCCA, RS - SpCCA adds to the random sampling rate r, which is an important parameter. In this section, we will briefly discuss the effect of sampling rate r on the recognition performance. Experiment chooses the fixed block size (as table 1) and random sampling rate changes from 10% to 100%, which takes 10% as the interval change. Figure 6 shows the effect of random sampling rate of the AR, Yale (6) (the first 6 samples of each kind are taken as training, and the others is taken as test) and ORL (6) and the others on the performance of RS - SpCCA. We can see from the figure: 1) For each test set, RS - SpCCA obtains the better performance than SpCCA in large scope (when r = 1, the RS-SpCCA converts to SpCCA). For Yale dataset, r changes from 0.2 to 0.9, and for ORL, r changes from 0.3 to 0.9. For the AR, it can get the optimal performance beyond at least five consecutive points. We have to mention that random sampling rate (or random sampling feature dimension) is a public problem that stochastic subspace method faces. Although we can't learn how to select the best random sampling rate from the figure, from the result, we can know that when random sampling rate r is set to 0.5 ~ 0.5, they all can obtain the optimal recognition accuracy than SpCCA. Therefore, RS-SpCCA still has operable.

The size of sub image is open issues faced by sub image method and it determines the performance of the sub image method. In the condition that random sampling rate is 0.5, we test the effect of different sizes of sub images in the Yale (6) and ORL (6) on recognition performance.

We can see the follows from the figure 7:

1) In addition to the specific size (2 x 2) , for the size of other sub image, RS-SpCCA based on random sampling obtains higher recognition performance than SpCCA;

2) Compared with SpCCA, the effect of the change of sub image size that RS-SpCCA suffers is smaller. On ORL and Yale dataset, the change ranges of SpCCA respectively are 63% and 21%, while RS - SpCCA only are 22% and 6.5%. Therefore, the RS - SpCCA is relatively more stable and less affected by the size of the sub image.

When the image is 2 x 2, RS-SpCCA makes 50% sampling and it can only get two pixels so that it is unable to reflect the local features of human face. Therefore, the effect naturally is not ideal. However, on the whole, RS-SpCCA always can get optimal performance than SpCCA in multiple points. Therefore, the RS - SpCCA has certain operability.

The computation efficiency of algorithm is an important indicator to measure the quality. In this paper, we propose RS - SpCCA method. According to make random sampling in each image set, construct more component classifiers. Therefore, compared with SpCCA and Semi-RS method, in training and testing phase RS-SpCCA has higher time complexity. Set original sample dimension is D and the training set is N and the number of sub image dimension is D (usually, D > > N > D). The number classifier of each image is K and random sampling rate is r (r < = 1) and the number of blocks is p (in this paper, D=d×p). The time complexity of key step

TABLE V. THE TIME COMPLEXITY OF THE ALGORITHM

	The stage of training				Testing phase			
	Global properties get (PCA)	Local characteristics of the access	Global and local information fusion	The overall time complexity	Relevant/sampling characteristics obtained	Similarity calculation	Voting decisions	The overall time complexity
SpCCA	$O(N^3)$	$O(D)$	$O(pN^3)$	$O(pN^3+D)$	$O(Dd)$	$O(DN)$	$O(P)$	$O(Dd+DN+p)$
RS-SpCCA	$O(N^3)$	$O(D)$	$O(KpN^3)$	$O(KpN^3+D)$	$O(KDdr^2)$	$O(KDNr)$	$O(Kp)$	$O(KDdr^2+KDNr+Kp)$
Semi-RS	-	$O(pd^3)$	-	$O(pd3)$	$O(KDdr^2)$	$O(KDNr^2)$	$O(Kp)$	$O(KDdr^2+KDNr+Kp)$

of the every algorithm and the overall time complexity are as shown in table 5. It is obvious that time complexity of RS-SpCCA in training phase is higher than SpCCA and Semi-RS. In the testing phase, the RS - SpCCA has exactly the same time complexity with Semi - RS (in fact, the RS - SpCCA and Semi - RS has exactly the same test process).However, algorithm of training process is done once, and it does not need to repeated operation. Therefore, here we will focus on considering the testing time. The test time of RS - SpCCA is mainly composed of the following parts: (a) related features acquire time; (b) similarity computation time of test sample and training sample; (c) the final classification result of the decision time. It can be seen from table 5 that for large-scale face recognition, the bottleneck of system speed lies in similarity calculation. Because RS - SpCCA construct more classifiers than SpCCA and it needs more times of similarity computation and comparison, in order to speed up the test, we use the classification strategy of "from coarse to fine" . That is to say, firstly, use global characteristics to build the classifier and according to the classifier, calculate the similarity of all samples in the test sample and training sample set and make list and retain a certain number samples (set its number is M) that have smaller difference with test samples. Then, test sample is limited in the M training samples of the local feature classifier. Due to the decrease of the times of similarity calculation, the test efficiency is improved. In theory, the bigger the M is, the more excellent the final performance is, but it consumes the more time. The smaller M is, less time it consumes, but performance will decrease. From the figure 8 (use AR77 as the test set) , we can see when the candidate image number M that is retained by the global classifier is less than 150, the recognition rate of system begins to drop significantly. Therefore, in order to ensure the accuracy of system, we select M = 150 as the ideal value. At the same time, we give the needed time of RS - SpCCA after optimization in the testing phase when M = 150 and make comparison with SpCCA method (Matlab 7.0, dual-core 2.0 G, 1 gb memory). As we can see from table 6, using mechanism "from coarse to fine" not only greatly increase the operation efficiency of RS - SpCCA, but also it has faster operation speed than SpCCA. Because under the ideal number of candidate sample, the recognition accuracy of RS-SpCCA can remain relatively unchanged, the mechanism “from coarse to fine” is effective. We should mention that test time which is as shown in table 6 is the overall test time in the condition of more than one test sample serial input. If use the parallel input form and use Matlab programming skills, it can reduce the overall test time. In addition, as a parallel multiple classifier system, our naturally use parallel processing to further reduce RS-SpCCA test time.

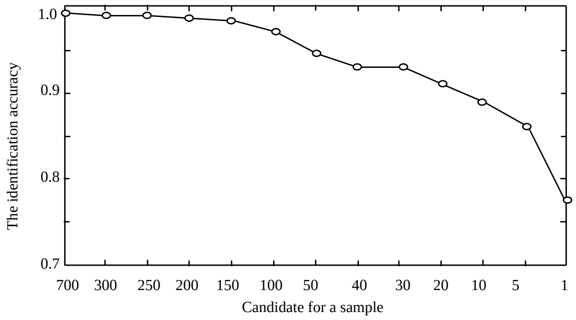


Figure 8. The effect of candidate sample M on recognition performance

TABLE VI. THE COMPARISON OF TEST TIME (S) OF DIFFERENT ALGORITHMS

	SpCCA	RS-SpCCA	From coarse to fine RS-SpCCA
AR77	1267	9648	483

IV. CONCLUSION

Based on CCA and the sub pattern methods, SpCCA is proposed in this paper and is compared with the DCV, pCA + CCA, SpPCA and Aw: SpPCA on the effect of face recognition. Experiments show that SpCCA global use CCA to fuse global features and local features, at the same time, it inherits the advantages of PCA method, such as the description ability of global information, the focus of sub pattern methods on local information, robustness of local change. It solves the problems that there exist small samples in the CCA, but it can show better recognition ability in face database, such as local illumination, expression and shelter and the other local changes. Three standard face database used in experiment is better than Semi-RS and SpCCA, which has better performance.

Although in this paper, RS - SpCCA uses PCA to extract global features and uses sampling feature set of original image feature as local features, in actual operation, it can use the other methods to extract global features (such as discrete Fourier transform) and replace PCA features or make appropriate processing for sampling features of sub image (for example, in order to make full use of the category information, use the LDA for feature extraction) so as to get better performance. In addition, the RS-SpCCA only uses local related features as discriminant features and discards the global features. At the same time, using CCA global features and local related features may further improve the performance of algorithm. In fact, the two groups of features can be replaced for any features that meet with dimension condition of CCA and PCA according to the necessity so as to look for better global and local features to improve the classification performance of the algorithm. The information that fuses the local and global information is improved for the information that fuses multiple levels information. Besides, the algorithm is extended to the field of regression, reconstruction, and others beyond classification, which prepares three directions of the research in the future.

REFERENCES

[1] D. Xu, Z. Y. Feng, Y. Z. Li, et al. Fair Channel allocation and power control for uplink and downlink cognitive radio networks. *IEEE, Workshop on mobile computing and emerging communication networks*, 2011 pp. 591-596

[2] W. Q. Yao, Y. Wang, T. Wang. Joint optimization for downlink resource allocation in cognitive radio cellular networks. *IEEE., 8th Annual IEEE consumer communications and networking conference,* 2011 pp. 664-668

[3] S. H. Tang, M. C. Chen, Y. S. Sun, et al. A spectral efficient and fair user-centric spectrum allocation approach for downlink transmissions. *IEEE., Globecom.,* 2011 pp. 1-6

[4] K. Ruttik, K. Koufos, R. Janttir. Model for computing aggregate interference from secondary cellular network in presence of correlated shadow fading. *IEEE, 22nd International symposium on personal, indoor and mobile radio communications*, 2011 pp. 433-437

[5] J. Naereddine, J. Riihijarvi, P. Mahonen. Transmit power control for secondary use in environments with correlated shadowing. *IEEE, ICC2011 Proceedings*, 2011 pp. 1-6

[6] Muhammad J. Mirza, Nadeem Anjum. Association of Moving Objects Across Visual Sensor Networks.*Journal of Multimedia*, Vol 7, No 1 (2012) , 2-8

[7] Haiping Huang, Hao Chen, Ruchuan Wang, Qian Mao, Renyuan Cheng.(t, n) Secret Sharing Scheme Based on Cylinder Model in Wireless Sensor Networks.*Journal of Networks*, Vol 7, No 7 (2012) pp. 1009-1016

[8] Xin Huang, Xiao Ma, Bangdao Chen, Andrew Markham, Qinghua Wang, Andrew William Roscoe.Human Interactive Secure ID Management in Body Sensor Networks. *Journal of Networks*, Vol 7, No 9 (2012) pp. 1400-1406

[9] Muhammad J. Mirza, Nadeem Anjum. Association of Moving Objects across Visual Sensor Networks. *Journal of Multimedia*, Vol 7, No 1 (2012) pp. 2-8

[10] Zhao Liangduan, Zhiyong Yuan, Xiangyun Liao, Weixin Si, Jianhui Zhao.3D Tracking and Positioning of Surgical Instruments in Virtual Surgery Simulation. *Journal of Multimedia*, Vol 6, No 6 (2011) pp. 502-509

[11] Kuo-Feng Huang, Shih-Jung Wu, Real-time-service-based Distributed Scheduling Scheme for IEEE 802.16j Networks. *Journal of Networks*, Vol 8, No 3 (2013) pp. 513-517

[12] Kasman Suhairi, Ford Lumban Gaol, The Measurement of Optimization Performance of Managed Service Division with ITIL Framework using Statistical Process Control. *Journal of Networks*, Vol 8, No 3 (2013) pp. 518-529

[13] Guang Yan, Zhu Yue-Fei, Gu Chun-Xiang, Fei Jin-long, He Xin-Zheng, A Framework for Automated Security Proof and its Application to OAEP. *Journal of Networks,* Vol 8, No 3 (2013) pp. 552-558

[14] R. Berangi, S. Saleem, M. Faulkner, et al. TDD cognitive radio femtocell network (CRFN) operation in FDD downlink spectrum. *IEEE, 22nd International Symposium on Personal, Indoor and Mobile Radio Communications,* 2011 pp. 482-486

[15] W. Ahmed, J. Gao, S. Saleem, et al. An access technique for secondary network in downlink channels. *IEEE, 22nd International Symposium on Personal, Indoor and Mobile Radio Communications*, 2011 pp. 423-427

[16] D. L. Sun, X. N. Zhu, Z. M. Zeng, et al. Downlink power control in cognitive femtocell networks. *IEEE., International conference on wireless communications and signal processing*, 2011 pp. 1-5

NIU Li-xia, her main research field is human resource management, Artificial intelligence.