A Weighted Regional Voting Based Ensemble of Multiple Classifiers for Face Recognition

Jing Cheng[1] and Liang Chen[2]

[1] Chengshi EDU, Vancouver, Canada
Jing.Cheng@alumni.unbc.ca
[2] University of Northern British Columbia Department of Computer Science 3333 University Way, Prince George, BC, Canada
lchen@ieee.org

Abstract. Face Recognition has become a heavily studied field of AI. Competing techniques have been proposed, both holistic and local, each has their own advantages and disadvantages. Recently, a unified methodology using a Regional Voting framework has improved the accuracy of all holistic algorithms significantly and is currently regarded as one of the best approaches. In this work, based on the success of regional voting, we developed a two layer voting system called Weighted Regional Voting Based Ensemble of Multiple Classifiers (WREC), which can embed all available face recognition algorithms. The first layer embeds a holistic algorithm into a Regional Voting framework. The second layer gathers the classification results of different algorithms from the first layer and then makes the final decision. Extensive experiments carried out on benchmark face databases show the proposed system is faster and more accurate than several other leading algorithms/approaches in every case.

1 Introduction

Face recognition is defined as a task to assign an identity to a face image. Broadly speaking, face recognition algorithms fall into two classes: local and holistic. Generally, most face recognition systems assume that the image is grayscale which is the same convention adopted in the approach outlined here where each pixel is represented by a positive number referring to the intensity of the gray.

1.1 Holistic Approaches

Holistic approaches treat the entire face image as a pattern to be classified where each image of h pixels tall and w pixels wide are represented by a slim vector of the length $h \times w$. Thus, each image is treated as a point in a high-dimensional space R^^{D}. Dimension D is the product of h and w.

Then, dimensional reduction is performed on the vectors: Given gallery images: G = $\{g_1, g_2, ..., g_n\}$ and the probe images Y, the projection matrix P is trained on the gallery images by solving a generalized matrix problem. The points having dimensions reduced are compared in a low-dimensional subspace using the nearest neighbour classification.

G. Bebis et al. (Eds.): ISVC 2014, Part II, LNCS 8888, pp. 482–491, 2014.

Holistic face recognition approaches based on statistical learning, such as the ones based on LDA, often suffer from the SSS (small-sample-size) problem, where the dimensionality of the sample images far exceeds the number of training sample images available for each subject [1], [2].

1.2 Local Approaches

Local approaches attempt to extract salient features/regions of the face. These regions are then used as classifiers, and the result of each region's classification is used to classify the overall image. A simple majority then decides the overall classification over all the regions.

However, the primary downside cited for holistic approaches is that it lacks the knowledge of the spatial structure of a face. The local approaches addressed this drawback while the pixels within the same region will affect the classification much more.

1.3 Regional Voting

A radical approach that embeds a holistic algorithm into a Regional Voting system suggested by Chen and Tokuda [3] has made some strides and leveraged the accuracy and stability against 'noise'. If only a few regions are contaminated by 'noise', then the effect of that noise is limited to only those regions. See Chen and Tokuda [4], [5] for an in-depth analysis of this stability. Its result has been regarded as the best in class.

Regional Voting successfully enhanced the performance of each holistic algorithm embedded. However, each algorithm is still working separately. Besides, it treats all regions equally important. This would suggest that applying a dynamic weighting distribution over the regions along with their effectiveness in recognizing a face and providing a mechanism for the approaches/algorithms to complement one another might be a rewarding avenue.

2 Proposed Algorithm

Here we present a two layered ensemble of multiple classifiers for face recognition called Weighted Regional Voting Based Ensemble of Multiple Classifiers (WREC) which can embed all available face recognition algorithms. It uses a voting scheme on classifiers' outcomes using weights based on facial regions' significance. Although the system presented here is embedding the holistic algorithms only, it could easily be extended to local approaches as well.

Based on the success of Regional Voting, the first layer of the system adds weight to different regions of the human face via a Local Weighted Voting framework. This idea exploits the fact that face regions are of different significance when recognizing a face. This concept can be traced back to the early 1970s in "Computer Recognition of Human Faces" [6]. The Local Binary Patterns by Timo Ahonen and others [7] shows another exploitation of the weight distribution.

The second layer enabled cooperation among different algorithms embedded via an ensemble of the multiple face recognition algorithms framework. This idea is motivated by the fact that different algorithms approach the face recognition challenge from different aspects. It would be a fruitful avenue to explore a new system, which has derived good points and qualities from the best already existing ones in face recognition literature.

2.1 First Layer: Local Weighted Voting Framework

First of all, we partition each image into $l \times m$ equally sized non-overlapping regions in a consistent manner. Thus, the relative locations of pixels have been taken into consideration. Then, within each region, perform classification independently.

Five of the top face recognition algorithms have been used here:
h_1: Principle Component Analysis (PCA)
h_2: Fisherface
h_3: Spectral Regression Dimension Analysis (SRDA) [8]
h_4: Spatially Smooth Version of Linear Discriminant Analysis (S-LDA) [9]
h_5: Spatially Smooth Version of Locality Preserving Projection (S-LPP) [10]

For each holistic algorithm $h \in H$, we implement a regional weighting based classification of any probe image during the test stage. Euclidean distance is used during the matching procedure for comparison purposes and the smallest distance value identifies the subject ID of the probe image.

Even after pre-processing, pupils, mouth, chin, forehead etc. may still end up located in different regions with respect to different face images. To avoid non-meaningful comparisons, shifting is implemented. We move each gallery region up to two pixels in four directions (north, south, west and east) to compensate for misalignment issues. All regional gallery images in all nearby positions (25 in total) are compared with the regional probe image using the nearest neighbour classifier and the results are stored. By selecting the closest one as the identity of the holistic algorithm h on that region, we get the classification on region r: $h^{\wedge r}\, p^{\wedge r}$. For each region, instead of contributing one vote for an identity $i \in I$, $w_{F(h,r)}$ is used as the ‘number of votes contributed (how the weighting of $w_{F(h,r)}$ is calculated, see Subsection 2.2). What the voting machine does is sum up the number of votes each identity gets. The one that gets the “biggest number” of votes is taken as the subject ID for probe image p by algorithm h.

At this point, for each region r, for each holistic algorithm h and by using all gallery images, we obtain a classifier $h^{\wedge r}\, p^{\wedge r}$, where p is a probe image and $p \in P$. The identity that wins the “biggest number” of votes is the final classification of holistic algorithm h on that probe image p. By now, the first layer of voting for classification is set up.

2.2 Weighting Scheme

Weights are based on facial regions' significance. Assuming that in our gallery there are N subjects $S_1, S_2, \ldots, S_N$. Each subject S_i has K images $G_{i1}, G_{i2}, \ldots, G_{ik}$. There is a set of holistic algorithms $H = \{h_1, h_2, \ldots, h_t\}$. For each holistic algorithm $h \in H$, on each region, we use a "leaving one out" strategy to test the effectiveness of the holistic algorithm on that region and take it as the weighting value of that region for algorithm h. For each j, $1 \leq j \leq K$, we select $G_{1j}, G_{2j}, \ldots, G_{Nj}$ as the testing set and take the remaining images in the gallery as the training set. By doing so, for each j, we find the number of correctly recognized images for each region by that algorithm: $r_{(h,j)}$. Dividing by N, $r_{(h,j)}/N$, we got the accuracy of that splitting[1]. In total, there are K splittings. The weighting value on region r for holistic algorithm h. $w_{(h,r)}$ is averaged over k times. It is formally defined in the following (see Equation 1):

$$w_{(h,r)} = \frac{\sum_{j=1}^{k} r_{(h,j)}}{N} . \tag{1}$$

Thus, $w_{(h,r)}$ stands for the average recognition accuracy on region r by holistic algorithm h and $0 \leq w_{(h,r)} \leq 1$. The regional weights generator compares the subspace regional feature vectors by Euclidean distance and selects the closest one as the classification sticking to the nearest neighbour classifier. After calculating the regional weighting of holistic algorithm h according to Equation (1), it implements one of the equations among Equations (2), (3) and (4).

Three different schemes are adopted for the final weight to be used during the test stage after accepting probe images. We call them "One Applies One", "One Applies All" and "Joint Weight" respectively. The following equations show the difference among the above three weighting schemes. In all cases, $w_{F(h,r)}$ stands for the final weight which is going to be assigned to the region r for holistic algorithm h.

One Applies One

$$w_{F(h,r)} = w_{(h,r)} . \tag{2}$$

One Applies All

$$w_{F(h,r)} = w_{(h_5,r)} . \tag{3}$$

Here, during the weighting evaluation on the training set, only the effectiveness of the S-LPP algorithm is tested. During the test stage having all probe images included, all algorithms (including S-LPP) use the weighting evaluated by S-LPP: $w_{(h5,r)}$.

[1] Splitting here refers to the division of gallery images into subTrain and subTest sets, unlike the one mentioned in Section. 3: Experiments, which refers to a component in the database.

Joint Weight

$$w_{F(h,r)} = \sum_{al=1}^{5} w_{(h_{al},r)} \,. \tag{4}$$

Here, $w_{(h_{al},r)}$ refers to the weighting assigned to region r evaluated by algorithm $h_{al} \in H$.

2.3 Second Layer

The second layer of voting is done among the different holistic algorithms. Each algorithm casts one vote to its identity and the final decision is made by a "winner takes all" strategy. That is, the identity voted by the majority of holistic algorithms is taken as the final result. After this round of voting, the final decision is made.

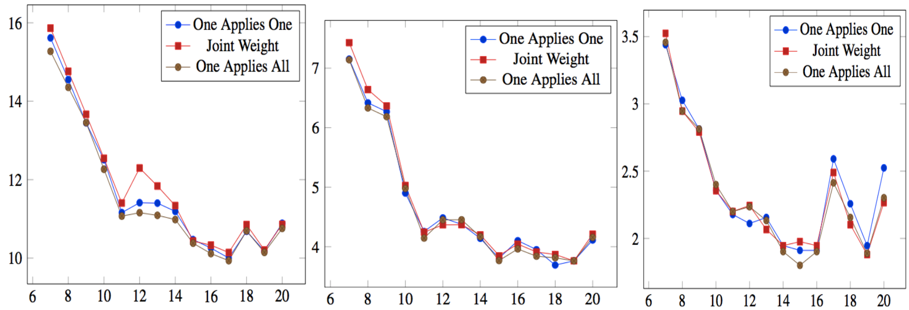
3 Experiments

In order to validate the WREC approach, three benchmark databases were used: the Yale Database [11], ORL Database [12] and Carnegie Melon University Pose, Illumination, and Expression database (CMU PIE) [13]. Following the custom of researchers in this field, the faces for all three databases were simply manually aligned by pupils and cropped to 64 × 64 pixels with 256 gray levels per pixel. To exclude any possible bias, including the pupil locating, rotating, scaling and cropping approach used for "standardization" during the pre-processing stage mentioned above, the UIUC versions of the ORL, Yale and PIE face databases are used. All face images are aligned based on pupil location and cropped to a common size. UIUC's version of the PIE database uses the near frontal poses (C05, C07, C09, C27, C29), which leaves us 11,554 face images. Each subject has 170 images, except subject 38, which has only 164 images. These pre-processed, scaled and rotated images were provided by Deng Cai[2].

Two baseline approaches (PCA and Fisherface) and three newly developed approaches (SRDA, S-LPP, and S-LDA) are compared. For the Eigenface approach, the eigenvector corresponding to the largest eigenvalue was removed accounting for noise caused by illumination. $\alpha = 0.01$ was selected for regularization in the SRDA and S-LPP approaches. For SLPP, cosine similarity was used to calculate the distances in the adjacency matrix. All gallery images were moved up to two pixels in each direction to make up for misalignment. All vectors (representation of images) were normalized to the length of 1. Probe images were linearly reduced and then classified according to nearest neighbour classification.

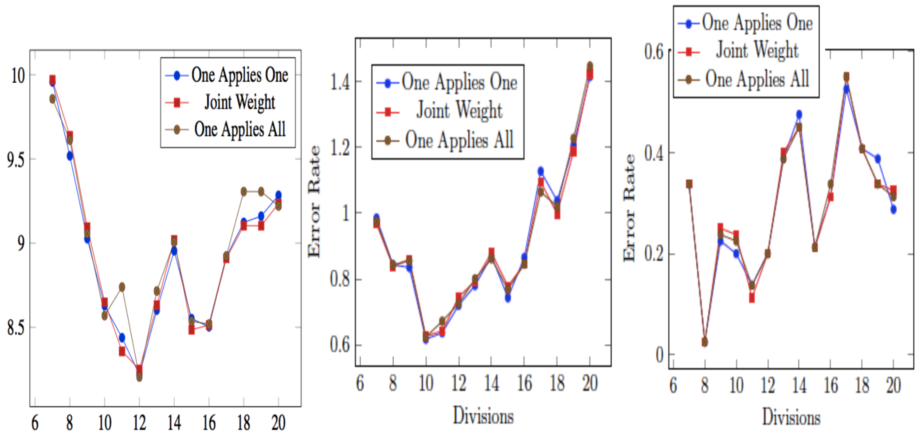
[2] All data and holistic algorithms were taken from: `http://www.zjucadcg.cn/dengcai/Data/FaceData.html`

WREC were carried out with different number of regions, from 7 × 7 to 20 × 20. That is, each image was first divided 7 times vertically, and 7 times horizontally, and then 8 times each direction and so on up to 20 horizontal and vertical divisions. This time, all three weighting schemes were implemented in this set of experiments. The results for the Yale and ORL database on 2, 5 and 8 training datasets are given in Fig. 1 and Fig. 2. (For all the figures in this paper, the 'x-axis' refers to the number of divisions and 'y-axis' refers to the error percentage)



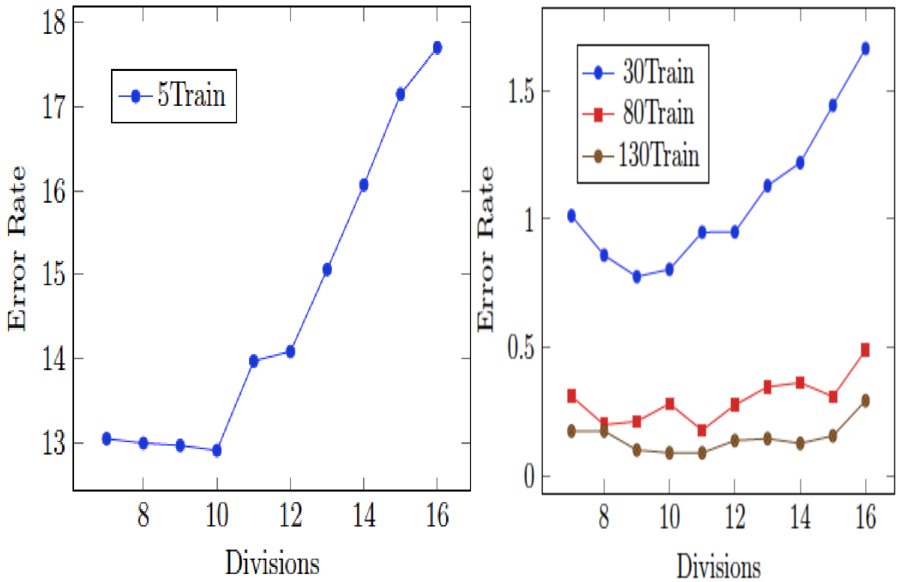
(a) 2 Training dataset (b) 5 Training dataset (c) 8 Training dataset

Fig. 1. WREC on Yale database on different training datasets with divisions from 7 up to 20



(a) 2 Training dataset (b) 5 Training dataset (c) 8 Training dataset

Fig. 2. WREC on ORL database on different training datasets with divisions from 7 up to 20



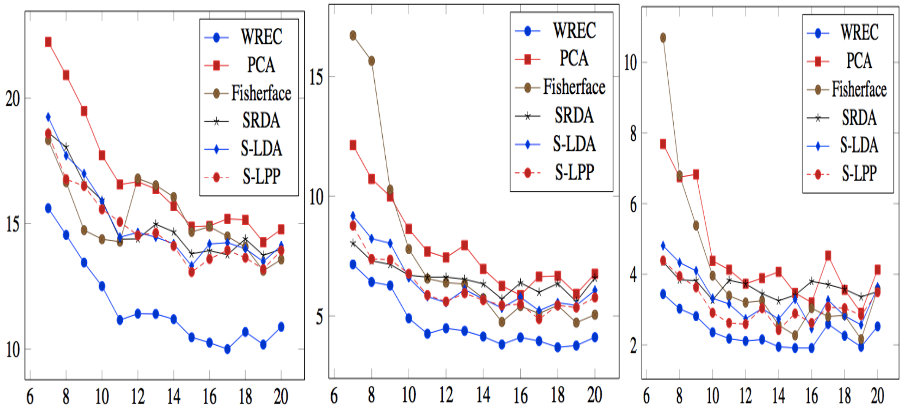
(a) PIE 5 Training dataset (b) PIE 30, 80 and 130 Training dataset

Fig. 3. WREC on PIE database on different training datasets with divisions from 7 up to 16

On the PIE database, due to time constraints as well as the possible distribution of the best results, the experiments were carried out on divisions from 7 up to 16 with only one weighting scheme: One Applies One. As the recognition accuracy for the 5 Train dataset differs a lot from the rest of the datasets, we put it aside in a separate figure to have a better representation of the result. The result is shown in Fig. 3.

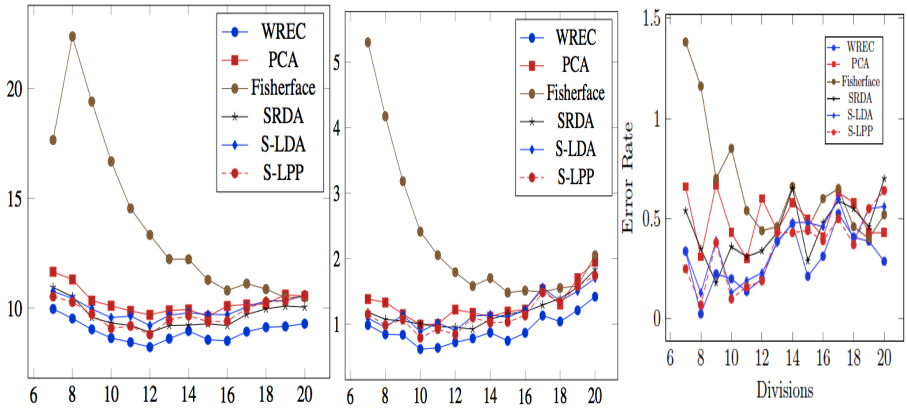
4 Analysis and Conclusion

The experiments demonstrate WREC's significant performance advantages compared to several other leading approaches. In a lot of cases, the error recognition rate drops more than half. Fig. 1, Fig. 2, and Fig. 3, all show the same pattern as Regional Voting [5] where the accuracy goes up as the number of regions increases. After a certain point, the accuracy begins to drop due to the regions becoming too small to distinguish from national voting. The above observations from the figures match precisely the theory of “Electoral College and Direct Popular Vote”.



(a) 2 Training dataset (b) 5 Training dataset (c) 8 Training dataset

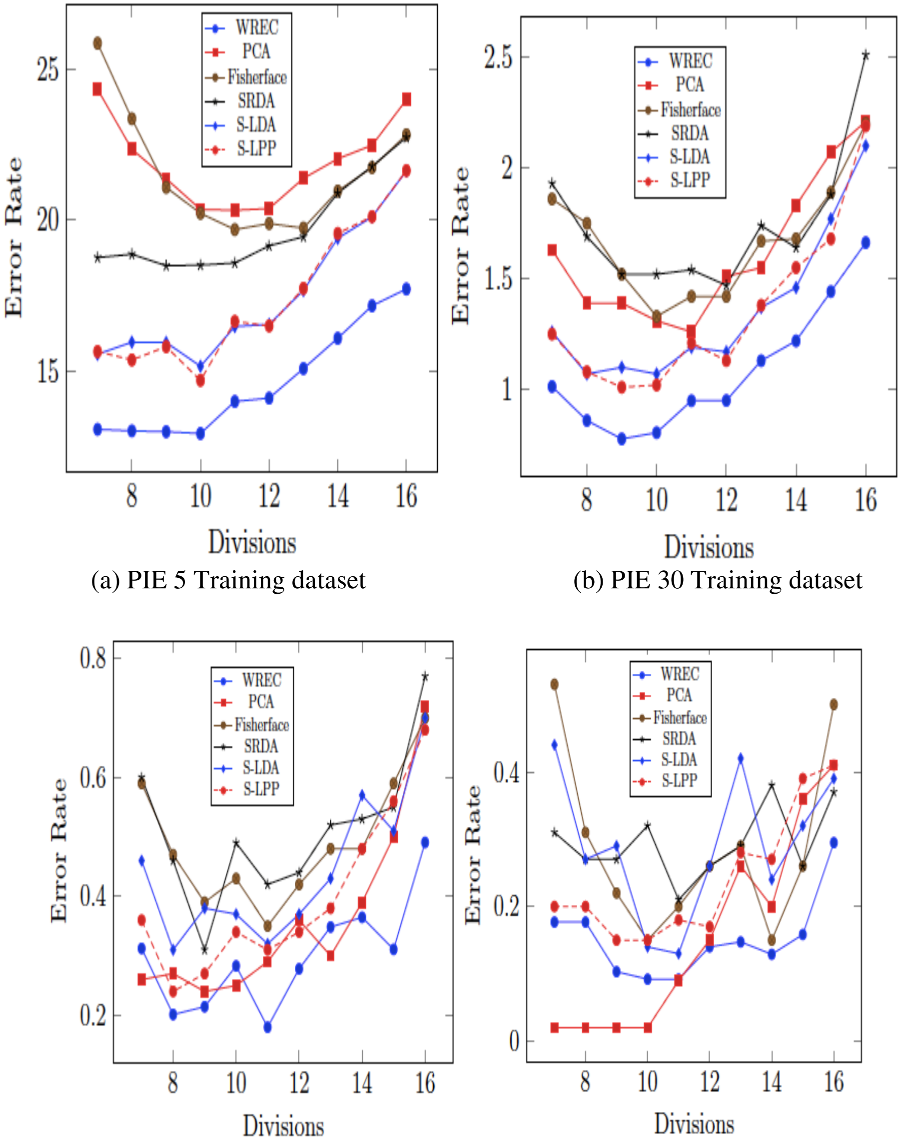
Fig. 4. WREC Compared to various individual holistic algorithms in different sized regions on Yale database



(a) 2 Training dataset (b) 5 Training dataset (c) 8 Training dataset

Fig. 5. WREC Compared to various individual holistic algorithms in different sized regions on ORL database

Besides, Fig. 4, Fig. 5 and Fig. 6 show the transition point (the division where the recognition performance peaks) in the WREC system appear earlier than Regional Voting. In these figures, the "One Applies One" weighting scheme is used for face recognition and the accuracy does not differ much from the multiple weighting schemes used for WREC. Suffice it to say, we wouldn't need to have all algorithms involved during weighting calculating stage.



(a) PIE 5 Training dataset (b) PIE 30 Training dataset

(c) PIE 80 Training dataset (d) PIE 130 Training dataset

Fig. 6. WREC Compared to various individual holistic algorithms in different sized regions on PIE database

"One Applies All" weighting scheme alone suffices almost all the time. Extensive experiments carried out on benchmark face databases show the proposed system's results holds a safe lead in every case over the already best in class results of Regional Voting. The same promising result on experiments of datasets with small number of images per person in the gallery images deserves emphasis as it belongs to an especially nasty problem: the SSS (small-sample-size) problem in the face recognition area, which has been mentioned earlier. Even an REC (a regional based ensemble of multiple classifiers) system shows promising results on datasets with a smaller number of gallery images.

References

1. Raudys, S., Jain, A.: Small sample size effects in statistical pattern recognition: Recommendations for practitioners. Pattern Anal. Mach. Intell. 13, 252–264 (1991)
2. Lu, J., Plataniotis, K.N., Venetsanopoulos, A.N., Li, S.Z.: Ensemble-based discriminant learning with boosting for face recognition. IEEE Transactions on Neural Networks 17, 178 (2006)
3. Chen, L., Tokuda, N.: A unified framework for improving the accuracy of all holistic face identification algorithms. Artif. Intell. Rev. 33, 107–122 (2010)
4. Chen, L., Tokuda, N.: Robustness of regional matching scheme over global matching scheme. Artificial Intelligence 144(1-2), 213–232 (2003)
5. Chen, L., Tokuda, N.: A general stability analysis on regional and national voting schemes against noise why is an electoral college more stable than a direct popular election? Artificial Intelligence 163, 47–66 (2005)
6. Kanade, T.: Computer recognition of human faces. Interdiscilplinary Systems Research 47 (1977)
7. Ahonen, T., Hadid, A., Pietikäinen, M.: Face recognition with local binary patterns. In: Pajdla, T., Matas, J. (eds.) ECCV 2004. LNCS, vol. 3021, pp. 469–481. Springer, Heidelberg (2004)
8. Cai, D., Member, S., He, X., Han, J., Member, S.: Srda: An efficient algorithm for large-scale discriminant analysis. IEEE Transactions on Knowledge and Data Engineering, 1–12 (2008)
9. Cai, D., He, X., Hu, Y., Han, J., Huang, T.: Learning a spatially smooth subspace for face recognition. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2007, pp. 1–7 (2007)
10. Xiaofei, H., Shuicheng, Y., Yuxiao, H., Partha, N., Hongjiang, Z.: Face recognition using laplacianfaces. IEEE Transactions on Pattern Analysis and Machine Intelligence 27, 328–340 (2005)
11. Yale Database, http://cvc.yale.edu/projects/yalefaces/yalefaces.html
12. ORL Database, http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html
13. CMU Database, http://www.ri.cmu.edu/research_project_detail.html?project_id=418&menu_id=261