

A face recognition optimization algorithm based on statistical characteristics of feature points

SUN Zhongyue¹, ZHOU Tianhui²

(1. Science and Technology Division, Beijing Jiaotong University, Beijing 100044, China;

2. College of Electrical and Information Engineering, Beijing University of Civil Engineering and Architecture, Beijing 100044, China)

Abstract: When calculating the similarity of human facial feature points, the elastic graph matching algorithm considers that each feature point contributes equally to face recognition, and hence is assigned to the same unit weight. Unfortunately, it is not usually true. This paper proposes a novel face recognition optimization algorithm based on the statistical characteristics of feature points. Based on the statistical theory and extracted Histograms of Oriented Gradients (HOG) characteristics, the feature points can be classified into the major and minor that assigned to different weights. The simulation results show that this proposed approach can provide a better performance behavior and a higher recognition rate using public available databases.

Key words: face recognition; elastic graph matching; HOG characteristic; statistical theory; weights optimization

基于特征点统计特征的人脸识别优化算法

孙中悦¹, 周天荟²

(1. 北京交通大学 科学技术处, 北京 100044; 2. 北京建筑大学 电气与信息工程学院, 北京 100044)

摘 要: 当计算人脸特征点相似度时, 弹性图匹配算法认为每个特征点对于人脸识别都起到相同的作用, 因此分配相同的单位权值。然而, 事实并非如此。基于特征点的统计特征, 提出一种新的人脸识别优化算法。利用统计理论和提取的 HOG 特征, 可以把特征点分成主要和次要两类, 并且赋予不同的权值。仿真结果表明所提出的方法不但性能较好, 而且识别率较高。

关键词: 人脸识别; 弹性图匹配算法; HOG 特征; 统计理论; 权值优化

中图分类号: TP391.41

文献标志码: A

1 Introduction

At present, face recognition has attracted much attention due to its features such as nature, intuition, non-contact, safety. It has become a hot problem in the field of pattern recognition and artificial intelligence. However, due to the complexity of facial structure, the diversity of facial expression and the

variability of face imaging process, etc, the automatic face recognition is still recognized as a challenging research field^[1-4].

Face image implies a lot of meaningful local and global face feature information. The information not only includes some direct features which can be perceived, e. g. eyes, nose, hair, but also contains some indirect features which can not be felt. These indirect

收稿日期: 2014-12-15

基金项目: 中央高校基本科研业务费专项资金资助(2014JBM172)

作者简介: 孙中悦(1974—), 女, 吉林长春人, 博士。研究方向为图像识别。email: sun_zhongyue@163.com.

features hide in the face images in some form. Generally, the major process of face recognition consists of two steps: ①extract relevant information from a given face image and express them in an effective way; ②compare the extracted information with the human face model library. It should be noted that the extracted information and the face model library should be described in the same way. For most of the face recognition methods, key points are located at the selection of the features which have larger difference among different individuals, and are stable for the same individual^[5-7]. Generally, face recognition algorithms for the stationary images can be classified into statistical feature and geometric feature based methods.

Statistical feature methods^[8-12] are to find a linear or nonlinear space transformation based on some performance indexes, and then project face images onto a dimensional reduced subspace where recognition are carried out. Their advantages lie in lower computational cost, stronger descriptive ability, better separable property, etc. However, a major problem is that they need a large number of training image data. The more little training data is, the more low recognition rate is. Utilizing a two layer structure, hidden Markov model^[13] can describe local and global face features efficiently to a certain extent. Unfortunately, the validity of the method is dependent on the extracted sample window features.

Geometric feature based methods^[14] try to recognize face using its facial components. Elastic Bunch Graph Matching (EBGM)^[15] is one of early presented methods. In EBGM, facial features are illustrated as topology graphs with several nodes at facial landmarks. Each node contains a set of Gabor wavelet coefficients that known as a jet. This method has stronger robustness in light and expression changes, and can realize accurate localization of facial feature points. Another evident advantage is that it needs less training sample images. It can be attained a better recognition effect with only one sample image. Nevertheless, EBGM has some shortcomings in the computational complexity and localization speed.

Most of the mentioned-above approaches consider that each feature point plays the same role in face

recognition. As we known, it is evident that this is coincided with statistical regularities. Following this research line, we present a new face recognition optimization algorithm based on the EBGM. The basic idea of the method is that we utilize a descriptor known as HOG^[16] to model facial landmarks. The HOG descriptor was initially presented by Lowe in his Scale Invariant Feature Transform (SIFT)^[17]. These features, which are extracted by SIFT, can be used for matching different views of an object or scene. Recently, the reader can see several improved HOG algorithms for face recognition^[18-19]. In the paper, utilizing statistics theory, we present a critical optimization algorithm for classifying facial feature points. Namely different feature points should be assigned to different weight values. This paper presents a comparison between a new face recognition optimization algorithm based on statistical characteristics of feature points and the original elastic graph matching algorithm based on HOG features. The experiments results show a better performance behavior using public available databases.

2 HOG based elastic graph matching algorithm

2.1 HOG descriptor

HOG descriptor is used to describe the appearance and shape of objects in an image pixel intensity gradients or edge of the directional distribution. The method is to divide the image into small units for grid connected region, and then collection box unit pixel gradient directions or edge direction histogram. Finally, combination of these histograms can constitute a feature descriptor. In order to improve the accuracy, we can use these local histogram normalized contrast images in a larger interval (block), calculate the density of each histogram in this interval (block), then according to the density values on the interval in each grid unit to do normalization. By normalizing, the image can obtain better stability to the change of illumination and shadow.

Compared with other feature description methods, HOG descriptor has many advantages. First, because the HOG approach is used to operate on the local grid unit of the image, so it can maintain a good

invariance to image geometry and optical deformation. Secondly, in the coarse spatial sampling, fine orientation sampling and strong local optical normalization conditions, some little tricks can be forgiven; these tiny movements can be ignored without affecting the detection results. Therefore, the direction of the gradient histogram method is particularly suitable for image of face recognition.

HOG descriptor is a local statistical data and the direction of the image gradient is around a critical point. As mentioned above, HOG descriptor is the last part of the SIFT algorithm. The algorithm is in some sense a more general method. It is invariant when changes the scale and rotation of the image. This is expressed only in the local minima of the scale space image feature point extraction and use of the image gradient information to find around the critical point the dominant direction. Both techniques have been proved very useful for images of arbitrary scaling and rotation. But when the image is not scaled or rotation facilitates identification, the information will be deleted.

In this paper, we assume that the exact location of the two eyes is known. Therefore, we do not want any scaling or rotation, the first two steps of the SIFT features will be ignored; we only consider the last, that is, HOG descriptor. Each HOG descriptor is a histogram. When the HOG feature value of the face image is extracted, each spatial bin is a 3×3 pixels square with the direction for the nine channels. The template $[-1, 0, 1]$ is applied to calculate the gradient of the nine channels and use trilinear interpolation weighting to obtain the gradient histogram of each unit. Then each block is normalized using the final output of a gradient histogram vectors as the HOG feature.

2.2 Algorithm realization

The main idea of elastic graph matching algorithm to recognize the new face image is based on the calibration of a series of feature points, and comparison similarity between the new image feature points and a set of training in the face database of each individual. Traditional elastic graph matching algorithm adopts a set of Gabor nodes as features to locate and match the feature points. In contrast, we select some

special points in the face as feature points such as nose, eyes, forehead, chin, etc. and then extract the HOG characteristic of each feature point, and similarity matching with the known feature points, ultimately used to detect the human face graphics.

Basically, the elastic graph matching algorithm can be decomposed into three steps: image normalization, create charts and the final image matching.

Image normalization step is to reduce the changes generated by the illumination changes, scaling and rotation. This paper uses the ORL and FERET face database. The normalization of the face are 92×112 and 80×80 pixels. The left and right eyes of ORL database face images are set at (30, 55) and (60, 55) pixel coordinates respectively, while the FERET database respectively (20, 65) and (60, 65). Graphics normalization is particular importance before computing the image gradients.

The next step is the detection of facial landmarks to create a facial figure. The successful recognition algorithms rely on a good facial landmarks selection: facial marks should be very unique between different people and easy to detect in a fully automated system. This paper using the facial figure is a project of the Central South University^[20] proposed the following 25 facial sign structure, as shown in Fig. 1.

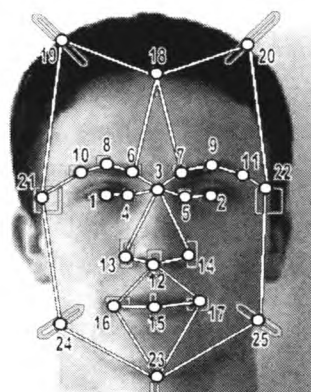


Fig. 1 Face graph

Finally is the image matching. In the process of recognition, we extract the characteristic values from the test set of image database by using the algorithm, by comparing the characteristics value of the image to determine the similarity of face images. The similarity formula is as follows,

$$\text{sim}(G, S) = \frac{1}{N} \sum_{i=1}^N \left(1 - \frac{g_i - s_i}{\max(g_i, s_i)} \right) \quad (1)$$

where, $sim(G, S)$ represents the similarity of face images; g_i and s_i respectively represent the two people face to be compared in the same feature points at the characteristic value; N is the number of facial feature points.

It is worth noting that not all areas of the face are contribution to face recognition equally. Studies have shown that the eyes and nose and their surrounding area are very important to identify, so more facial landmarks are on there. Because each feature point is different from the role of the recognition result, we assign different feature points to different weights which could improve the recognition rate of face images.

2.3 Modeling and locating the facial landmarks

In the algorithm, each face is described by a face graph (FG); the facial figure is composed of strategically located facial landmarks and their corresponding descriptors

$$FG = \{X_i, J_i = HOG(X_i), 1 \leq i \leq 25\} \quad (2)$$

where, X_i represents the serial number of feature points and J_i represents the corresponding HOG value. When creating a new FG, we need a set of models to automatically locate facial landmarks. These sets of models should account for changes of expression, hair styles, illumination, etc. We build a new model for each landmark using a set of HOG descriptors extracted from a set of training images. Let $fbg_i(k)$ be the HOG descriptor of the i -th keypoint from the k -th training face (each FG has 25 landmarks). We define the face bunch graph (FBG) as: $fbg_i(k) (1 \leq i \leq 25, 1 \leq k \leq N_f)$, where, N_f is the number of training sample.

The location of the descriptors $fbg_i(k)$ on a training image is manually marked. However, facial landmarks in test images are automatically located. This is done in an iterative process in which a new landmark uses the information of previously detected landmarks to reduce the search area. The process to detect the i -th ($i > 2$) facial landmark.

Step 1 Initial estimation of the facial landmark location. X_{si} is based on the mean of displacements between the i -th keypoint and the j -th ($j < i$) keypoint using the information of the geometry of the graphs in the FBG.

Step 2 Calculate the HOG descriptor on the previous location $HOG(X_{si})$.

Step 3 Compare $HOG(X_{si})$ with all HOG descriptors of the models in fbg_i .

$$K_{\min} = \min(HOG(X_{si}) - fbg_i(k)) \quad (3)$$

Step 4 Define a search area S_i around X_{si} . The extent of the search area depends on the particular keypoint as shown in Fig. 1. We empirically set the search areas considering the dispersion of the location of facial landmarks in each $fbg_i(k)$ of the FBG.

Step 5 Refine the initial estimation of the i -th keypoint using the descriptor $fbg_i(K_{\min})$.

$$X_i = \min_{X \in S_i} (HOG(X) - fbg_i(K_{\min})) \quad (4)$$

3 Algorithm optimization

3.1 Optimization Theory

In the traditional algorithm, each individual face images include 25 feature points. The algorithm considers that the recognition results for each feature point are the same, thus the same weight of all the feature points are given. But experimental tests show that the results are not the case. For different faces, the results of the contribution of each feature point are not the same. Some feature points on the results of the impact are bigger, and some feature points affect very small on the results. According to the statistical characteristics of the human face, to give different feature points and different weights could greatly improve the face recognition rate.

When feature matching, if the increase in the weights of a feature point can improve the recognition rate of the experiment, it illustrates the influence of the feature point on the recognition result is larger and is denoted as the main characteristic; if the right value of a feature point increased could not improve the recognition rate, it proves that the feature point is less impact on the recognition result and is denoted as the secondary feature point.

Increase the weight of characteristic values is to increase the similarity of feature points. Calculate the degree of similarity is to calculate the difference. The similarity is greater which leads to the smaller difference. In the process of the calculation, if the weights of the main feature points are increased, that is multiplied by a constant greater than 1 and is equivalent to

enlarge the differences between the feature points which lead to the final results bigger than any other characteristic points. Therefore, the optimized similarity formula is as follows,

$$sim(G, S) = \frac{1}{N} \sum_{i=1}^N \omega_i \left(1 - \frac{g_i - s_i}{\max(g_i, s_i)} \right) \quad (5)$$

where, ω_i represents the value of the corresponding weights of each feature point. If the feature points are as the main feature points, ω_i should take a constant greater than 1. The specific value of ω_i is discussed in the next section.

3.2 Weights selection

First, the main feature points and the secondary feature points are extracted. Specific methods are adopted based on the tested multiple groups of face images. According to the statistical characteristics, more than one person face graph groups are selected and a group of feature points are extracted from the group of each face map. Then the multiple sets of feature points are compared and one of the highest frequency of several is selected as the final main feature point. As these points are more than the statistical results of the map groups, so has the universality, it can be applied to other diagram groups.

After the key feature points are identified, the weights of the main features of the point are selected. If ω_i is too small, it is not enough to significantly affect the recognition rate. Under normal circumstances, greater weights are better. However, the extracted feature points are the statistical results and not all diagram groups are the main features of the point. In this case, the greater weights may offset by the effect of the other major feature point optimization. According to the analysis, ω_i gradually increases from 1 and the recognition rate is gradually increasing. When ω_i is greater than a certain value, the recognition rate has negative impacts and it begins to decline. The turning point of ω_i is the optimal weight.

4 Experimental results

This paper selects two face databases in the experiment, the British ORL (Olivetti Research Laboratory) face database^[21] and FERET (Face Recognition Technology) face database^[22].

4.1 ORL based experiments

40 images are selected in the experiment containing 10 different expressions of human face. 40 face object extraction of an arbitrary subset constitutes a group. In the verification, an image is chosen from each object in the map group used as a target face, and then the other images of the object are selected as a testing face. The proposed recognition algorithm is used to compare the similarity between face images of each person which could get the recognition rate.

The main feature points are selected as below: 10 different images subset assuming that a feature point in the 10 map group as the number of occurrences of the main feature points over 7 times. The point is considered for the main features of the feature points. Tab.1 shows the main feature points for each map group in the experiment (1 to 25 describe 25 feature points of human faces respectively).

Tab.1 Each map group feature point statistics based on ORL

Map series	Main feature points
Group 1	1, 3, 4, 5, 6, 9, 11, 12, 13, 14, 15, 17, 18, 20, 22
Group 2	5, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 23
Group 3	3, 4, 5, 8, 9, 11, 12, 14, 16, 17, 18, 19, 20, 21, 23
Group 4	1, 4, 5, 7, 9, 11, 12, 14, 15, 16, 17, 18, 20, 22, 24, 25
Group 5	2, 3, 4, 7, 9, 11, 12, 13, 15, 16, 18, 19, 21, 23
Group 6	1, 2, 3, 5, 7, 9, 11, 12, 14, 15, 17, 19, 20, 22, 25
Group 7	1, 3, 5, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 20, 23
Group 8	1, 2, 3, 5, 9, 10, 11, 13, 15, 16, 17, 19, 20, 21, 24
Group 9	1, 2, 5, 6, 7, 9, 11, 13, 14, 15, 16, 17, 18, 19, 22, 23
Group 10	1, 4, 5, 6, 9, 12, 13, 15, 16, 18, 19, 20, 21, 22, 25

According to statistical theory, the frequencies of ORL face graph feature points as the main features are shown in Fig. 2.

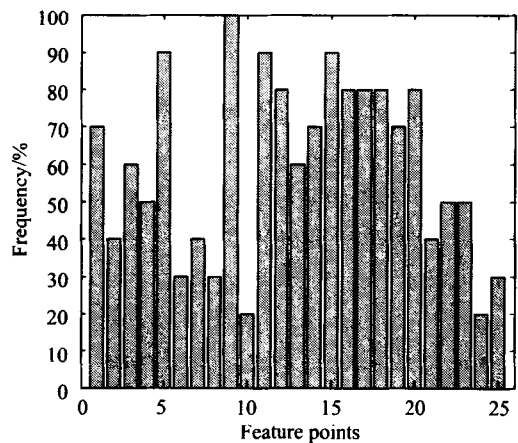


Fig.2 Frequencies of ORL face map feature points as the main features

Fig. 2 shows that the main features of the face map in ORL face database is 1, 5, 9, 11, 12, 14, 15, 16, 17, 18, 19, 20.

In order to verify the main selected features and their choice of weights, 5 new sets of maps are selected, testing the recognition rate when ω_i value is 1, 2, 3, 4, 5, 6, respectively. The results are shown in Fig. 3.

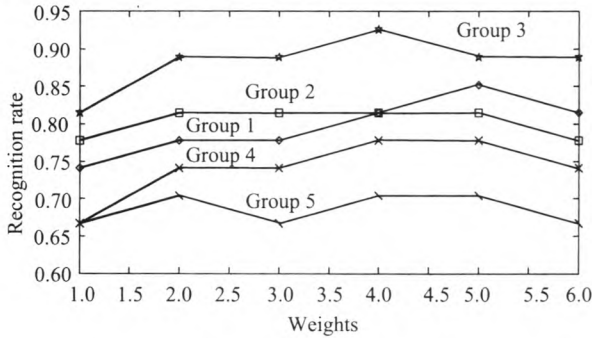


Fig. 3 Recognition rates comparison under different weights based on the ORL face database

4.2 FERET based experiments

40 images are selected containing 7 different expressions in the FERET face database. Any subset of the 40 individual face objects are extracted to form a map group. The feature points of face database selected map group in the experiment are shown in Tab. 2.

Tab. 2 Each map group feature point statistics based on FERET

Map series	Main feature points
Group 1	1, 3, 5, 7, 9, 11, 12, 14, 16, 18, 19, 24, 25
Group 2	1, 2, 3, 6, 9, 11, 12, 14, 15, 16, 17, 20, 22, 25
Group 3	4, 5, 8, 9, 10, 11, 13, 15, 16, 17, 18, 19, 20, 24, 25
Group 4	1, 3, 4, 5, 9, 12, 13, 15, 16, 17, 19, 20, 21, 23, 24
Group 5	1, 2, 5, 7, 9, 11, 12, 14, 15, 17, 18, 19, 21, 22, 24, 25
Group 6	1, 3, 6, 8, 10, 12, 13, 15, 16, 17, 18, 19, 20, 21, 24, 25
Group 7	3, 4, 6, 9, 11, 12, 13, 14, 15, 16, 17, 19, 20, 23, 25
Group 8	1, 2, 7, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19, 20, 23, 24
Group 9	1, 3, 4, 7, 9, 11, 12, 14, 15, 16, 19, 21, 24, 25
Group 10	1, 3, 5, 8, 9, 11, 12, 14, 15, 16, 17, 18, 20, 25

Similarly, according to statistical theory, the frequencies of FERET face graph feature points as the main features are shown in Fig. 4.

Fig. 4 shows the main features of the face map in FERET face database is: 1, 5, 9, 11, 12, 14, 15, 16, 17, 18, 19, 20. Similarly, 5 new groups are selected to test the recognition rate when ω_i are assigned to different values. The experimental results are shown in Fig. 5.

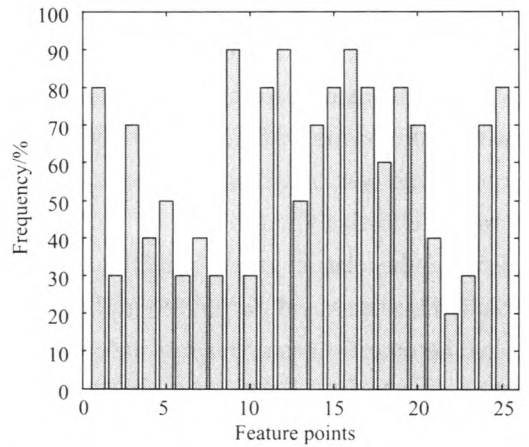


Fig. 4 Frequency of FERET face map feature points as the main features

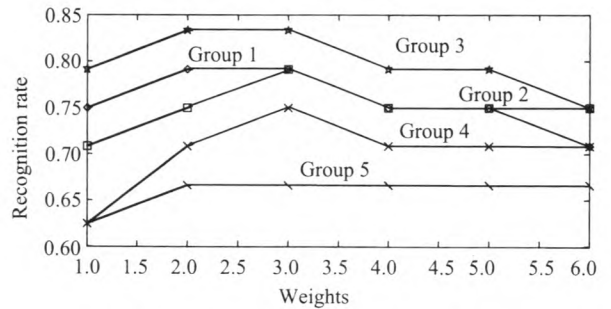


Fig. 5 Recognition rates comparison under different weights based on the FERET face database

4.3 Summary

In the experiment, $\omega_i = 1$ is the recognition rate of EBGm algorithm without optimization. It can be seen from Fig. 3 and Fig. 5, the recognition rate has improved when optimizing the weights of the feature points for different face groups. It improves by 4% to 10% in ORL database while 4% – 8% in FERET. In addition, when ω_i gradually increases from 1, the recognition rate for each face group overall is upward trend; when ω_i increases to a certain value, the recognition rate begins to decrease. Different group curves in the figures have different inflection points. However, they are concentrated in a certain interval, for example, ORL database diagram group is concentrated in 4 to 5 and the FERET in 2 – 3. Therefore, the value of a constant is selected in this interval as a major feature point weight.

5 Conclusions and Prospects

This paper proposes a new elastic graph matching face recognition optimization method to realize the automatic calibration of the facial feature points com-

bined with statistical theory. The HOG features of the facial feature points are extracted base on categories. Then the main feature points are selected and different feature points with different weights are given to obtain the best matching. Moreover the recognition rate is effectively improved.

Because the elastic graph matching algorithms have high robustness and better recognition effects, they have become the mainstream of face recognition algorithms and derive a lot of improvements and optimization methods to further improve the performances. But there also exists some problems and needs further research, for example, the computational complexity of this algorithm is high, the operation time is long and the speed of recognition is slow. In this regard, we can combine this algorithm with other algorithms in the premise of keeping good recognition rates to achieve the feature vector dimensions reduction.

References:

- [1] Chellappa R, Wilson C L, Sirohey S. Human and machine recognition of faces: a survey [J]. *Proceedings of the IEEE*, 1995, 83(5): 705 – 741.
- [2] Samal A, Iyengar P A. Automatic recognition and analysis of human faces and facial expressions: a survey [J]. *Pattern Recognition*, 1992, 25(1): 65 – 77.
- [3] Zhao W, Chellappa R, Phillips P J, et al. Face recognition: a literature survey [J]. *ACM Computing Surveys*, 2003, 35(4): 399 – 459.
- [4] Phillips P J. Improving face recognition technology [J]. *Computer*, 2011, 44(3): 84 – 86.
- [5] Pmiz-del-Solar J, Navarrete P. Eigenspace-based face recognition: a comparative study of different approaches [J]. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, 2005, 35(3): 315 – 325.
- [6] Wu J, Zhou D. Survey of face recognition [J]. *Application Research of Computers*, 2009, 26(9): 3205 – 3209. (in Chinese)
- [7] Geng Y. On the face recognition technology and its applications [J]. *Friend of Science Amateurs*, 2011(4): 131 – 133. (in Chinese)
- [8] Poon B, Amin M A, Yan H. Performance evaluation and comparison of PCA based human face recognition methods for distorted images [J]. *International Journal of Machine Learning and Cybernetics*, 2011, 2(4): 245 – 259.
- [9] Sun J, Zhang H, Sun C. A new method of face recognition based on singular value decomposition [J]. *Science Technology and Engineering*, 2010(25): 6204 – 6208. (in Chinese)
- [10] Zhang J, Xiao D. Face recognition method based on multi-scale adaptive LDA [J]. *Computer Engineering and Design*, 2012, 33(1): 332 – 335. (in Chinese)
- [11] Kwak K C, Pedrycz W. Face recognition using an enhanced independent component analysis approach [J]. *IEEE Transactions on Neural Networks*, 2007, 18(2): 530 – 541.
- [12] Gao H. Research and simulation of face recognition based on NMF and its improvements [J]. *Computer Simulation*, 2011, 28(12): 232 – 235. (in Chinese)
- [13] Chien J, Liao C. Maximum confidence hidden Markov modeling for face recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, 30(4): 606 – 616.
- [14] Li J, Lei L, Li H, et al. Dynamic face recognition based on geometric features [J]. *Science Technology and Engineering*, 2010, 28(10): 6894 – 6898. (in Chinese)
- [15] Wiskott L, Fellous J M, Kruger N, et al. Face recognition by elastic bunch graph matching [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(7): 775 – 779.
- [16] Bay H, Tuytelaars T, Gool L V. Surf: speeded up robust features [C] // *Proceedings of the 9th European Conference on Computer Vision*, 2006: 404 – 417.
- [17] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision Archive*, 2004, 60(2): 91 – 110.
- [18] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C] // *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005, 1: 886 – 893.
- [19] Mikolajczyk K, Schmid C. A performance evaluation of local descriptors [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(10): 1615 – 1630.
- [20] Bolme D S, Beveridge J R, Teixeira M, et al. The CSU face identification evaluation system: its purpose, features and structure [C] // *Proceedings of the 3rd International Conference on Computer Vision Systems*, 2003: 304 – 313.
- [21] AT&T Laboratories Cambridge. The ORL database of faces [DB/OL]. (1999-03-02). <http://www.cam-orl.co.uk/facedatabase.html>.
- [22] Phillips P J, Moon H, Rauss P. The FERET evaluation methodology for face-recognition algorithms [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(10): 1090 – 1104.