Research Article

A Face Recognition Method Based on Multifeature Fusion

Shengxi Ye ![ORCID]

Wuhan Shiny Technology Co., Ltd, Wuhan 430205, China

Correspondence should be addressed to Shengxi Ye; 2015223030063@stu.scu.edu.cn

Face recognition is widely used in daily life and has an important supporting role for social management. Face recognition is mainly based on historical accumulation data to confirm people's identities in unknown samples and obtain valuable intelligence information. For the problem of face recognition, this paper proposes a multifeature joint adaptive weighting algorithm framework. In this method, a number of different types of features are first used to describe the face characteristics. The selected features should be as complementary as possible, and the overlap redundant information should be reduced to the greatest extent, so as to ensure the performance and efficiency of multifeature fusion. In the classification stage, based on the joint sparse representation model, the multiple types of features are characterized, and their reconstruction error vectors for the corresponding features of the test sample are calculated. The joint sparse representation model can examine the correlation between different types of features, thereby improving the accuracy of representation and fully integrating the advantages of multiple types of features. At the same time, in view of the simple superposition of reconstruction errors in the traditional sparse representation model, this paper uses a random weight matrix to comprehensively consider the weighted reconstruction errors under different weight conditions, so as to obtain statistical decision quantities for the final decision. The framework proposed in this paper can adapt to different multifeature combinations and has good practicability. In the experiment, training and test sets are constructed based on public face image data sets to test the proposed method. The experimental results show that the method in this paper is more effective and robust compared with some present methods for face recognition.

1. Introduction

In the recent COVID-19 epidemic, the face recognition system has played an important role in quickly confirming the identities of people and improving the effectiveness of joint prevention and control. Face recognition is a widely studied problem in pattern recognition [1–4]. Face recognition is a traditional supervised classification problem, in which a reliable classifier is obtained under the training of some labeled samples, and the samples of unknown categories are confirmed. Under the framework of pattern recognition, this paper focuses on the research of face image recognition, that is, confirming the people's identities based on the acquisition of face images. Judging from the existing literature, the general pattern recognition methods can generally be divided into two stages: feature extraction and classification. Feature extraction is used to obtain various features of the objects in face images, such as principal component analysis (PCA), Zernike moments of the target area, Gabor texture, and histogram of oriented gradients (HOGs), so as to achieve the goal of describing high-dimensional images with low-dimensional feature vectors [5–8]. In the classification stage, it is mainly to select a suitable classifier to process the result of feature extraction and obtain its corresponding object category. Common classifiers include K nearest neighbors (KNN), support vector machines (SVM), and sparse representation-based classification (SRC) [9–13]. In recent years, deep learning algorithms have become an important tool in remote sensing image target detection and recognition and have been widely used and verified [14–18].

In traditional face recognition methods, a single category of features is usually used in the feature extraction stage. In fact, the scenes acquired by face images are relatively complex, and the characteristics of a single type of feature are often difficult to fully describe the characteristics of the object. In addition, the adaptive recognition scenes are very limited. For this reason, researchers have developed the multifeature decision fusion for the problem of face recognition

[18–20]. Judging from the reported results, the joint application of multiple features can effectively improve the performance of face recognition, which has advantages over a single category of features [21–25]. However, the current multifeature fusion method has certain shortcomings. One is that the relevance of multicategory features is not fully utilized. For the different types of characteristics of the same object, there are mutual differences (complementarity) between them, but also a certain correlation. For example, in each type of feature space, the correlation between the features of the test samples and the features of different training samples tends to be the same. Therefore, making full use of this inherent correlation can further utilize the advantages of various features. Second, in the decision-making combination of different features, the same weight is usually used, that is, the importance of different features to the recognition result is considered to be the same. In fact, in different recognition scenarios, the advantages of various features are often different, and methods with equal weights have certain disadvantages. Considering the above shortcomings, this paper proposes a multifeature joint adaptive weighting decision fusion framework for the problem of face recognition. Aiming at the problem of insufficient relevance of multiple types of features, this paper uses joint sparse representation [17, 18] to represent the multiple types of features of face images, so as to use their relevance to improve the reconstruction accuracy. For the problem of equal weighting, this paper adapts to the reconstruction error of different feature outputs, by constructing a random weight matrix, using multiple sets of different weight vectors for weighted fusion. Considering fusion results of a large number of weight vectors, this paper constructs statistics based on the mean and variance and constructs the final decision variables based on them, which is used to identify different types of objects in face images. The experiment further constructs training and test samples containing multiple types of targets based on the public face image data sets. Then, the proposed method is tested and compared with some existing face recognition methods. The experimental results reflect the effectiveness and robustness of this method.

2. Multifeature Joint Representation

The traditional sparse representation model is mainly for a single input. On this basis, a sparse representation classification algorithm is derived, and it has been widely used and verified for target recognition problems such as faces and vehicles. The joint sparse representation model can be considered as a multitask form of sparse representation, with the ability to simultaneously represent multiple inputs [17, 18]. At the same time, through appropriate correlation constraints, the joint representation accuracy of multiple inputs can be effectively improved. Specifically in this article, the joint sparse representation is employed to analyze and process multiple features of face images. Assuming that the K different features of the test sample y are $y^{(1)}$, $y^{(2)}$, ..., $y^{(K)}$, they are sparsely represented as follows:

$$\min_{\beta} \left\{ g(\beta) = \sum_{k=1}^{K} \left\| y^{(k)} - A^{(k)} \alpha^{(k)} \right\| \right\}. \tag{1}$$

In equation (1), $A^{(k)}$ is the dictionary corresponding to the kth feature, which is generally constructed by the feature vectors from the training samples; $\alpha^{(k)}$ is the corresponding sparse representation coefficient vector; $\beta = [\alpha^{(1)} \, \alpha^{(2)} \cdots \alpha^{(K)}]$ is the coefficient matrix, which contains the sparse representation coefficient vectors of different features.

Although the above optimization goals achieve formal unity, they are not substantially different from solving each sparse representation problem independently. For this reason, the joint sparse representation model is modified based on equation (1) and appropriately constrains the coefficient matrix β to construct a new objective function as follows:

$$\min_{\beta} g(\beta) + \lambda \|\beta\|_{2,1}. \tag{2}$$

Under the constraint of the ℓ_1/ℓ_2 norm, the sparse coefficient vectors of different types of features tend to have similar distribution rules, thereby reflecting the correlation between them and improving the overall representation accuracy.

Based on the solved sparse coefficient matrix, the joint sparse representation model makes the object label decision according to the minimum error criterion in equation (3).

$$\text{identity}(y) = \min_{i} \sum_{k=1}^{K} \left\| y^{(k)} - A_i^{(k)} \alpha_i^{(k)} \right\|, \tag{3}$$

where $A_i^{(k)}$ and $\alpha_i^{(k)}$ are the local dictionary and coefficient vector corresponding to the ith class in the kth feature space, respectively.

The decision criterion in equation (3) actually believes that different types of features have the same importance for the final recognition result, so it is only a simple error accumulation. In the application of actual face recognition, due to changes in acquisition conditions and changes in the environment, different types of features may have different importance. Therefore, the traditional method of error summation with equal weights has certain limitations. It is necessary to find adaptive weights to take advantage of different types of features.

3. Decision Fusion

As mentioned above, the decision of the traditional joint sparse representation model can be considered as the result of equal weight fusion, and the characteristics and advantages of different inputs are not fully utilized. To this end, this paper uses multiple sets of random weight vectors for decision fusion, which together form a random weight matrix W as follows:

$$W = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1N} \\ w_{21} & w_{22} & \cdots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{K1} & w_{K2} & \cdots & w_{KN} \end{bmatrix}, \tag{4}$$

where each column in the matrix corresponds to a random weight vector, and the elements satisfy the following constraints:

$$\sum_{k=1}^{K} w_{ki} = 1, w_{ki} \geq 0. \tag{5}$$

Assuming that the error obtained from the reconstruction of the kth feature of the type using the ith local dictionary of the type is r_k^i, the weighting process of a set of weight vectors in the random matrix is described as follows:

$$R_n^i = \begin{bmatrix} r_1^i & r_2^i & \cdots & r_K^i \end{bmatrix} \begin{bmatrix} w_{n1} \\ w_{n2} \\ \vdots \\ w_{nK} \end{bmatrix}. \tag{6}$$

For a random weight matrix containing N random weight vectors, the ith class can obtain corresponding results to form a fusion error vector $R = [R_1^i \ R_2^i \ \cdots \ R_N^i]$.

Each constraint in the random weight vector has strong randomness under the condition of satisfying equation (5). With the continuous increase in the scale of the weight vector, the fusion error vector can be well reflected in the dominance of different features. Therefore, the relevant statistical value of the fusion result can better reflect the decision result of multitype feature fusion. The decision variable constructed in this paper for face recognition is as follows:

$$J = m + \lambda\sigma. \tag{7}$$

In equation (7), m and σ are the mean and variance of a certain category of fusion error vector, respectively; λ is the adjustment parameter. Correspondingly, for C different training categories, the corresponding C decision variable values can be calculated accordingly, which are recorded as $J_1, J_2, \cdots, J_C$. According to the definition of the decision variable, when it is smaller, the reconstruction error is smaller and tends to be stable. Therefore, the probability of the test sample belonging to this category is greater at this time. Therefore, this article compares the values of decision variables in different categories and considers that the category with the smallest value is the target category of the test sample. Based on the above discussion, the key steps of the method in this paper are shown in Figure 1, which can be specifically described as follows. First, multitype feature extraction is performed on all test samples, and a corresponding dictionary is constructed. Secondly, the corresponding multitype characteristics of the test sample are obtained. Then, the joint sparse representation is used to calculate the reconstruction errors of different categories for various features. Finally, a random weight matrix is constructed, and the decision variables corresponding to each category are calculated to determine the target category of the test sample.

4. Experimental Results

4.1. Data Sets. In this paper, the Yale-B face image data set is used for the performance evaluation of the proposed method. The data set includes 45 facial images from 10 people, each of which is acquired under different lighting conditions. The image size is 32 pixels × 32 pixels. Figure 2 shows some exemplar samples from the data set. In the experiment, 20 images of each person are randomly selected for training, and 25 samples are left as test samples.

This paper focuses on designing a framework for multifeature fusion recognition of face images and does not clarify the types and forms of multifeatures. For this reason, in the experiment process, it is mainly compared with the existing multifeature (classifier) fusion method. Specifically, the first comparison method is the joint sparse representation model (denoted as JSR). This method is consistent with the general idea of this article, but after the joint sparse representation, random weight fusion and statistical analysis decision-making are not performed. The second comparison method is based on the D-S evidence theory (denoted as DS) [15]. The decision results of different features (classifiers) are fused based on D-S. The third method uses a voting mechanism (denoted as voting) [16] for different features (classifiers). The result of the decision is processed. For these methods, this paper selects three typical faces based on the existing literature, namely, Zernike moment, HOG feature, and Gabor feature. These three types of features can, respectively, reflect the target geometry, local texture, and spectral characteristics, so they have a certain degree of complementarity.

4.2. Results and Discussion. The experiment is first carried out based on the original test samples. Table 1 compares the test results of several decision fusion methods on the original samples. It can be seen that the average recognition rate of the method in this paper is the highest, reflecting its performance advantages. Compared with the method of directly using the joint sparse representation model, the method in this paper increases the average recognition rate by 2.31% through further random weight fusion and statistical analysis. The improvement effect is significant. Compared with the D-S fusion and voting method, the fusion mechanism in this paper has obtained a higher recognition rate, reflecting its advantages. Table 1 also compares the average time consumptions of different methods for a single test sample. Due to the introduction of random weight processing, the efficiency of this method is slightly lower than that of the joint sparse representation method. Since the weighting process of random weights is a linear operation, the efficiency reduction is not obvious. The other two types of decision fusion methods involve the combined use of multiple features and classifiers, so the time cost is significantly increased. In summary, the method in this paper has strong advantages for face recognition problems.

The actual collection process of the face image may be interfered by different degrees of noises, resulting in a low signal-to-noise ratio (SNR) of the test sample to be identified. This experiment mainly investigates the recognition performance of the proposed method under early noise
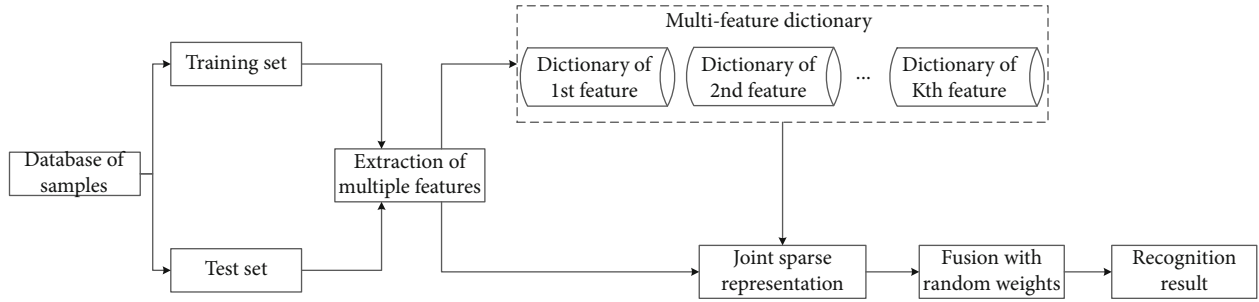
FIGURE 1: Basic flow of the proposed method in this paper.

FIGURE 2: Examples of face images in Yale-B data set.

TABLE 1: Recognition performance of different methods on original samples.

	Method type			
	Proposed	JSR	D-S	Voting
Average recognition rate (%)	96.8	94.5	95.7	95.1
Average time consumption (ms)	79.5	76.3	92.4	89.6

TABLE 2: Performance of different methods on noisy samples (%).

	SNR				
	-10	-5	0	5	10
Proposed	72.1	79.2	85.1	89.9	95.4
JSR	66.5	73.6	77.8	84.6	92.3
D-S	67.0	74.1	78.4	84.8	92.4
Voting	68.1	76.3	81.4	86.6	93.4

interference conditions. Specifically, first, the noise simulation is performed on the original test samples according to the relevant ideas in the literature [14] to obtain test sample sets with different noise levels. Then, the proposed method and the comparison method are used to identify test samples with different SNRs. Their results under different noise levels are shown in Table 2. From the overall trend, all the methods are affected by noise interference to varying degrees, and there is a significant performance degradation. In contrast, the proposed method can maintain performance advantages under different noises and shows stronger noise robustness. With random weights and statistical analysis, the advantages of various features for noise interference can be fully utilized. At the same time, the optimization processing of the joint sparse representation solution process also reduces the influence of noise to a certain extent, thereby comprehensively promoting the robustness of the proposed method against noise interference.

5. Conclusion

This paper proposes a multifeature joint adaptive weighting framework for the design of face image recognition problems. The method uses multiple complementary features to describe the target characteristics, so as to merge their advantages as much as possible. The joint sparse representation is used as the basic classifier to jointly characterize multiple features and use their relevance to improve the accuracy of the representation. For the reconstruction error vectors output by different types of features, a random weight matrix is used for weighted fusion and statistical analysis. Finally, a decision is made based on the statistics of the results under multiple different weights, and a reliable recognition result is obtained. In the experiment, the proposed method is tested based on the public face image data sets. The recognition performance of the method in this paper is better than that of some existing methods for the

original samples and noise samples. The results verify the effectiveness of the proposed method. The method in this paper constructs a framework for multifeature decision fusion, which is suitable for the use of different feature combinations. In the follow-up, the research will conduct in-depth research on the selection of multiple features, give quantitative constraint criteria, and further improve the performance of the recognition method.

Data Availability

The data set can be accessed upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] R. He, J. Cao, L. Song, Z. Sun, and T. Tan, "Adversarial cross-spectral face completion for NIR-VIS face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 5, pp. 1025–1037, 2020.

[2] R. He, X. Wu, Z. Sun, and T. Tan, "Wasserstein CNN: learning invariant features for NIR-VIS face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 7, pp. 1761–1773, 2019.

[3] F. Liu, Q. Zhao, X. Liu, and D. Zeng, "Joint face alignment and 3D face reconstruction with application to face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 3, pp. 664–678, 2020.

[4] X. Geng, Z.-H. Zhou, and K. Smith-Miles, "Individual stable space: an approach to face recognition under uncontrolled conditions," *IEEE Transactions on Neural Networks*, vol. 19, no. 8, pp. 1354–1368, 2008.

[5] J. Lu, V. E. Liong, X. Zhou, and J. Zhou, "Learning compact binary face descriptor for face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 10, pp. 2041–2056, 2015.

[6] C. Fu, X. Wu, Y. Hu, H. Huang, and R. He, "DVG-face: dual variational generation for heterogeneous face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 6, pp. 2938–2952, 2022.

[7] C. Ding, C. Xu, and D. Tao, "Multi-task pose-invariant face recognition," *IEEE Transactions on Image Processing*, vol. 24, no. 3, pp. 980–993, 2015.

[8] A. R. Rivera, J. R. Castillo, and O. O. Chae, "Local directional number pattern for face analysis: face and expression recognition," *IEEE Transactions on Image Processing*, vol. 22, no. 5, pp. 1740–1752, 2013.

[9] M. Luo, J. Cao, X. Ma, X. Zhang, and R. He, "FA-GAN: face augmentation GAN for deformation-invariant face recognition," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 2341–2355, 2021.

[10] R. Weng, J. Lu, and Y.-P. Tan, "Robust point set matching for partial face recognition," *IEEE Transactions on Image Processing*, vol. 25, no. 3, pp. 1163–1176, 2016.

[11] Z. Li, D. Gong, Y. Qiao, and D. Tao, "Common feature discriminant analysis for matching infrared face images to optical face images," *IEEE Transactions on Image Processing*, vol. 23, no. 6, pp. 2436–2445, 2014.

[12] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: application to face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037–2041, 2006.

[13] W. Zhang, S. Shan, X. Chen, and W. Gao, "Local Gabor binary patterns based on Kullback–Leibler divergence for partially occluded face recognition," *IEEE Signal Processing Letters*, vol. 14, no. 11, pp. 875–878, 2007.

[14] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, 2009.

[15] L. Fang and S. Li, "Face recognition by exploiting local Gabor features with multitask adaptive sparse representation," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 10, pp. 2605–2615, 2015.

[16] G. Hermosilla, J. L. Verdugo, G. Farias, E. Vera, F. Pizarro, and M. Machuca, "Face recognition and drunk classification using infrared face images," *Journal of Sensors*, vol. 2018, Article ID 5813514, 8 pages, 2018.

[17] D. R. Kisku and S. Rana, "Multithread face recognition in cloud," *Journal of Sensors*, vol. 2016, Article ID 2575904, 21 pages, 2016.

[18] M. Sajid, N. Ali, S. H. Dar et al., "Data augmentation-assisted makeup-invariant face recognition," *Mathematical Problems in Engineering*, vol. 2018, Article ID 2850632, 10 pages, 2018.

[19] W.-S. Chen, B. Pan, B. Fang, M. Li, and J. Tang, "Incremental nonnegative matrix factorization for face recognition," *Mathematical Problems in Engineering*, vol. 2008, Article ID 410674, 17 pages, 2008.

[20] G. Sang, J. Li, and Q. Zhao, "Pose-invariant face recognition via RGB-D images," *Computational Intelligence and Neuroscience*, vol. 2016, Article ID 3563758, 9 pages, 2016.

[21] K. Ko, H. Gwak, N. Thoummala, H. Kwon, and S. H. Kim, "SqueezeFace: integrative face recognition methods with LiDAR sensors," *Journal of Sensors*, vol. 2021, Article ID 4312245, 8 pages, 2021.

[22] T.-X. Jiang, T.-Z. Huang, X.-L. Zhao, and T.-H. Ma, "Patch-based principal component analysis for face recognition," *Computational Intelligence and Neuroscience*, vol. 2017, Article ID 5317850, 9 pages, 2017.

[23] T. Li and Z. Zhang, "Robust face recognition via block sparse Bayesian learning," *Mathematical Problems in Engineering*, vol. 2013, Article ID 695976, 13 pages, 2013.

[24] Y. Sun, Z. Ren, and W. Zheng, "Research on face recognition algorithm based on image processing," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9224203, 11 pages, 2022.

[25] D. Song, J. Luo, C. Zi, and H. Tian, "3D face recognition using anthropometric and curvelet features fusion," *Journal of Sensors*, vol. 2016, Article ID 6859364, 8 pages, 2016.