



Improved image representation and sparse representation for face recognition

Xuqin Wei¹ · Yun Shi¹ · Weiyin Gong¹ · Yanyun Guan¹

Received: 9 September 2021 / Revised: 13 January 2022 / Accepted: 11 May 2022 /

Published online: 2 June 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Sparse representation is of great significance to the research of face recognition. Due to factors such as illumination, angle, and facial features, different face images of the same subject have great differences under different conditions, which will cause great difficulties in image classification. To solve this problem, this paper proposes a novel image classification algorithm. The algorithm uses an improved image representation method to generate virtual samples that not only better preserve the large-scale information and global features of the original training samples, but can also be regarded as another representation of objects. Combining multiple representations of images can effectively improve the accuracy of image classification. Moreover, this paper designs a simple and efficient weight fusion scheme to fuse the original training samples and virtual samples and obtain the final classification distance of the test sample. Experimental results on multiple face databases show that the proposed algorithm has higher classification accuracy than other state-of-the-art algorithms.

Keywords Sparse representation · Face recognition · Image representation · Weight fusion

1 Introduction

In the field of computer vision and pattern recognition, face recognition has always received extensive attention [4, 9, 17, 33]. However, face recognition still faces challenges due to the

✉ Xuqin Wei
xuqinwei@126.com

Yun Shi
shiyunzhou2004@163.com

Weiyin Gong
Gw-Yin@hotmail.com

Yanyun Guan
guanyanyun2021@126.com

¹ College of Mathematics and Computer Science, Liupanshui Normal University, Liupanshui, China

influence of illumination, angle, and facial expressions, as well as the small number of samples. Sparse representation based classification (SRC) [24] has excellent feature extraction and classification recognition performance. After its application in face recognition, many important research results have been obtained [2, 3, 7, 12, 30, 31]. Inspired by SRC, Zhang et al. [36] proposed a more general model, Collaborative representation based classification (CRC), which not only has lower complexity but also achieved very competitive results. Keinert et al. [10] proposed a Group-Sparse Representation-based method, which is robust to noise, occlusions, and disguises. Considering the pose, expression, and misalignment changes of face images, MA et al. [16] used a sparse representation method with an adaptive weighted spatial pyramid structure. In view of the small number of training samples, the fast kernel sparse representation algorithm [5] classifies the original data by nonlinear mapping to the high-dimensional data space, which has high operating efficiency. Considering that real face recognition systems require high execution efficiency, Ye et al. [34] improved an extended sparse representation classification algorithm that is suitable for many single training samples. The JLSRC algorithm [25] reveals the potential relationship of the data by learning two low-rank reconstruction dictionaries and embeds non-negative constraint and an Elastic Net regularization in the coefficient vector of the dictionary to ensure that the dictionary has the better discriminative ability and is more robust to noise. To better deal with the structural noise of face images such as occlusion, illumination, etc., Tan et al. [22] used a novel geometric sparse representation model with a single image, which better retains the original geometric structure information of the image.

For sparse representation, the dictionary learned from training samples has an important influence on the result of image classification. Dictionary learning is an important branch of sparse representation. Due to its excellent performance, it has become an important method for studying face recognition [13, 18]. Classical dictionary learning methods include K-SVD [1], LC-KSVD [8], etc. Considering the importance of label constraints to the discriminative ability of the dictionary, Zhang [35] proposed a dictionary learning method using the labels of training samples, while Shrivastava et al. [21] only used part of the labeled data to obtain the dictionary. Li et al. [11] used both label constraints and local constraints to construct dictionary discriminants and obtained very competitive results in face recognition experiments. Due to the small number of samples, Lin et al. [14] used training samples of different facial expressions to generate a dictionary. Zhang [38] et al. proposed a sample-expanded multi-resolution dictionary learning method, which not only increases the diversity of image representation but is also more suitable for the classification of images with different resolutions in reality. Xu et al. [32] observe the object through two different angles of the image row and column.

To reduce the influence of external conditions such as illumination and angle, and improve the accuracy of image classification, this paper proposes a novel image classification algorithm. First, the algorithm uses a new image representation method to generate virtual images. This not only ensures that the large-scale information of the original training sample is retained in the virtual sample but also reduces the difference between different face images of the same person. Based on sparse representation, the algorithm uses original training samples and virtual samples to linearly represent test samples and obtains the classification scores corresponding to the two types of samples. Then, this paper designs a simple and efficient weight fusion scheme. This scheme can convert the classification scores of the original image and the virtual image into the final classification distance. Experimental results on multiple face databases show that the proposed algorithm has great advantages in classification accuracy.

The other parts of the paper are organized as follows. Section 2 introduces sparse representation and related works. Section 3 describes the algorithm steps and related implementation details. Section 4 gives the principle and analysis of the proposed algorithm. Section 5 shows the experimental results and analysis. Section 6 provides the conclusions of the paper.

2 Related works

This section mainly introduces the basic principles of sparse representation and related works. We define some notations that will be used in this paper.

Assuming that the subjects of the data set have C classes, and the number of training samples of each class is n , then the original training samples are recorded as $X = (X_1, X_2, \dots, X_C)$, where $X_i = (x_{i1}, x_{i2}, \dots, x_{in})$. X_i represents the training samples of the i -th class, and the column vector x_{ij} represents the j -th training sample of the i -th class.

The basic principle of sparse representation theory is to use a set of over-complete bases to linearly represent the input signal, and the obtained linear combination coefficients can approximately represent the input signal with a certain degree of sparsity. The most classic algorithm is the sparse representation based classification (SRC) [24]. The model of this algorithm can be expressed as:

$$\min_{\alpha} \|y - X\alpha\|_2^2 + \lambda \|\alpha\|_1 \quad (1)$$

Where y represents the test sample, which can also be considered as the input signal. X represents the set of training samples, α is the coefficient vector, λ is the parameter and $\|\alpha\|_1$ represents the $L1$ norm constraint.

The computational complexity of solving the sparse representation model is high, which is difficult to meet the requirements of practical applications. Therefore, convert the $L1$ norm to $L2$ norm constraint as follows:

$$\min_{\alpha} \|y - X\alpha\|_2^2 + \lambda \|\alpha\|_2 \quad (2)$$

This model is also called Collaborative representation based classification (CRC) and has higher computational efficiency. Then the difference between the test sample y and its estimated value $X_i\alpha_i$ can be calculated to achieve classification, as follows:

$$Identify(y) = \arg \min_i \|y - X_i\alpha_i\|_2^2 \quad i = 1, 2, \dots, C \quad (3)$$

At present, many algorithms based on sparse representation have been developed and achieved excellent performance in face recognition. For example, a two-phase test sample sparse representation method can accurately classify test samples [26]. In the first stage, the method represents the test sample as a linear combination of all training samples and determines the nearest M neighbors of the test sample. In the second stage, these neighbors are used to linearly represent the test sample and the result is used for classification. Furthermore, Using the symmetrical image [27] or mirror image [28] of the face image can not only increase the number of samples but also effectively improve the performance of image classification.

Xu et al. [30] used a novel image representation method to obtain virtual images and fused the original images and virtual images by a weighted fusion method to obtain the final classification results of test samples. On this basis, Zheng et al. [39] extended two new image transformation methods and achieved good performance.

3 Proposed algorithm

3.1 Algorithm steps

This section introduces the implementation process of the proposed algorithm. The specific steps of the algorithm are as follows.

Step 1. All samples of the data set are divided into two parts: training set and test set.

Step 2. Use formula (11) to generate virtual images. All images are converted to virtual images.

Step 3. Solve the coefficients that use training samples to linearly represent test samples.

According to the idea of sparse representation, the training sample X can be used to linearly represent the test sample. Assuming that y is a test sample, y can be linearly combined by X , as follows.

$$y = X\alpha = \sum_{i=1}^{Cn} x_i \alpha_i \quad (4)$$

x_i is the i -th column of X , representing the i -th training sample, and α_i is the coefficient vector of x_i . According to the regularized least square method to solve the equation, the linear combination coefficient α of y is as follows:

$$\alpha = (XX^T + \lambda E)^{-1} X^T y \quad (5)$$

λ is a constant and E is the identity matrix.

In the same way, the virtual sample V is used to linearly represent y_v as follows:

$$y_v = V\beta = \sum_{j=1}^{Cn} v_j \beta_j \quad (6)$$

y_v is a virtual sample generated by the test sample y using equation (11). v_j is the j -th column of V , representing the j -th virtual sample, and β_j is the coefficient vector of v_j . Then the linear combination coefficient β of y_v is as follows:

$$\beta = (VV^T + \lambda E)^{-1} V^T y_v \quad (7)$$

Step 4. Calculate the classification distance.

The distance between the test sample y and the original training sample of the i -th subject is recorded as d_1^i .

$$d_1^i = \|y - X_i \alpha_i\|_2 \quad i = 1, 2, \dots, C \quad (8)$$

Similarly, the distance between y and the virtual training sample of the i -th subject is recorded as d_2^i .

$$d_2^i = \|y_v - V_i \beta_i\|_2 \quad i = 1, 2, \dots, C \quad (9)$$

Step 5. Weight fusion scheme.

Use formula (14) to fuse the original image and the virtual image to obtain the classification distance d_i of the test sample.

Step 6. Classify the test sample.

d_i reflects the distance between the test sample and each subject, and the class of its minimum value is the class of the test sample. The test sample is classified by the following methods:

$$s_j = \arg \min_i d_i \tag{10}$$

The test sample is classified into the class of the minimum value s_j of d_i , that is, the j -th class.

3.2 Generating of virtual images

The proposed algorithm uses a novel image representation method to convert the original training samples into virtual images. The conversion process is as follows.

$$V = \sqrt[4]{X \cdot (255 - X)} \tag{11}$$

V represents the virtual image, and X is the original training sample. It can be easily seen that V and X have the same structure, that is, $V = (V_1, V_2, \dots, V_C)$, where $V_i = (v_{i1}, v_{i2}, \dots, v_{in})$. V_i represents the virtual image generated by the original training sample of the i -th subject, and the column vector v_{ij} is the virtual image generated by the j -th training sample of the i -th subject.

This image representation method is based on the conversion of pixel values of gray images. From the above equation, it can be seen that the pixels with the pixel value of p or $(255 - p)$ in the original image are all mapped to the same pixel value, which reflects the symmetry of the virtual image. In addition, Xu et al. [30] argue that medium pixels have stability and are more suitable for the classification of deformable objects such as faces. Equation (11) allows the pixel values of the virtual image to be better distributed around pixels of medium intensity and significantly reduces the differences between pixels. This transformation result can better preserve the large-scale information and global features of the original image, which are often beneficial for image classification [39].

3.3 Weight fusion scheme

The image classification algorithm is applied to the original image and the virtual image respectively to obtain the distances between them and the test sample, denoted as d_1^i and $d_2^i (i = 1, 2, \dots, C)$.

The distance represents the contribution when the training sample set linearly represents the test sample, the smaller the distance, the greater the contribution. Therefore, the algorithm sorts the distance between the test sample and the original face image and the virtual face image in ascending order. Assume that the sorting result of d_1^i is $q_1^1, q_1^2, \dots, q_1^C$, and the sorting result of d_2^i is $q_2^1, q_2^2, \dots, q_2^C$. Since small distances such as $q_1^1, q_1^2, q_2^1, q_2^2$ will have an important impact on the classification results, the weight fusion scheme pays more attention to these small distances.

The fusion weights of the original face image and the virtual face image are denoted as w_1 and w_2 , respectively.

$$w_1 = \frac{q_1^1 + q_1^2}{(q_1^1 + q_1^2) + (q_2^1 + q_2^2)} \tag{12}$$

$$w_2 = 1 - w_1 \tag{13}$$

The calculation method of the distance between the test sample and the i -th subject is as follows:

$$d_i = w_1 d_1^i + w_2 d_2^i \quad i = 1, 2, \dots, C \tag{14}$$

4 Rationale of the proposed algorithm

Image representation is a very promising method in the field of computer vision and pattern recognition. The appropriate image representation plays an important role in the final result of image classification. For deformable objects, such as human faces, using image representation to generate potential samples can reflect the changes of the original samples to a certain extent. Especially in the case of insufficient samples, generating virtual samples is very helpful to improve the accuracy of image classification.

This paper proposes a novel image representation method to generate virtual images, implemented as (11). In the virtual image, the difference between pixels is significantly reduced. To a certain extent, for the same subject, the difference between its different images has been reduced, while the virtual image retains the rich large-scale information of the original image. Figure 1 shows the first face image of the first subject in the ORL face database. Figure 2 shows the distribution of pixel values of the image. Figure 3 shows the distribution of the pixel values of the virtual image generated by the image. Figure 4 shows that the pixel values of the virtual image are symmetrical.

From Figs. 2, 3, and 4, it is obvious that the pixel value of the virtual image is significantly reduced, and the distribution of the pixel value is more concentrated. After the two pixels with larger pixel value differences in the original image are mapped to the virtual image, the difference is significantly reduced. When the pixel value of the original image varies from 0 to 255, the pixel value of the virtual image has symmetry.

Figure 5 shows the normalized data distribution of the original face image and its corresponding virtual sample. It can be seen from the figure that the proposed image representation has a low correlation with the original face image, which indicates that the virtual image can be used as a supplement to the original image to represent the face. In other words, the virtual sample is also another representation method of the subject.

Figure 6 shows the original gray face images of a subject in the Georgia Tech face database and the virtual face images generated by them. It can be seen from the figure that

Fig. 1 The first face image of the first subject in the ORL database



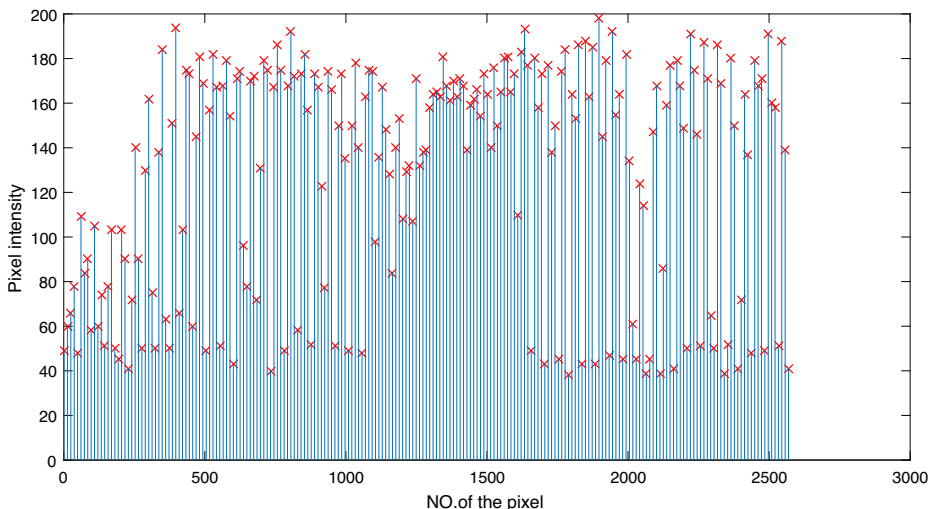


Fig. 2 Pixel value of the first face image of the first subject in the ORL database

the virtual face image still has complete facial features. Therefore, the original training sample and the virtual sample can be combined to represent the face object. Besides, the weight fusion scheme is a very simple and efficient method, which is often applied to image classification. This paper designs a new weight fusion scheme to fuse the classification scores of the original face image and the virtual image. This scheme can automatically obtain the fusion weight of the original face image and the virtual image and has high computational efficiency.

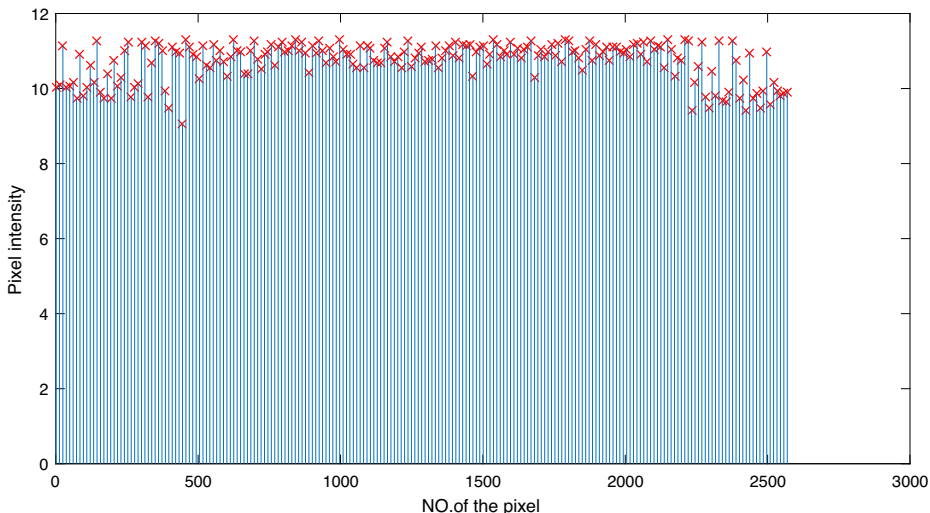


Fig. 3 The pixel value of the virtual image generated by the first face image of the first object in the ORL database

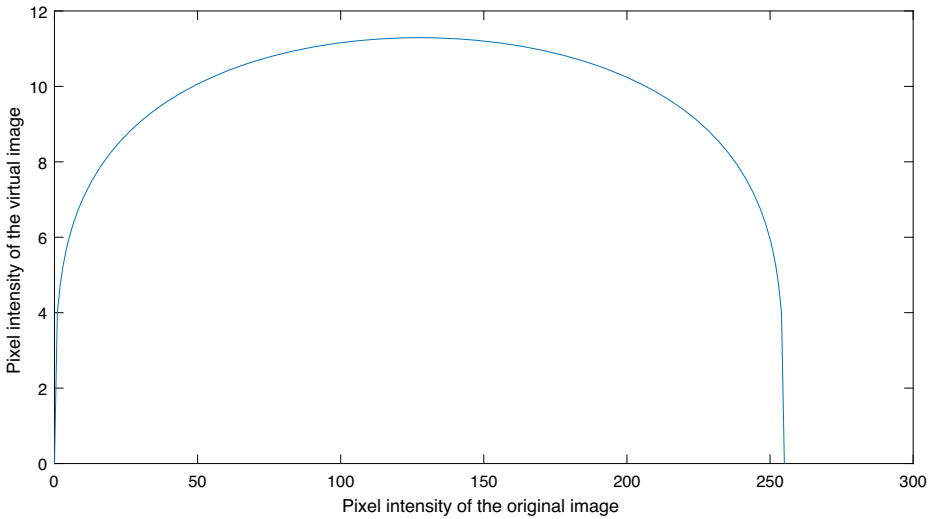


Fig. 4 Symmetry of pixel values of the virtual image

5 Experimental and results

To reflect the superiority of the proposed algorithm, we conducted experiments on the ORL face database [20], Georgia Tech face database [6], and FERET face database [29]. The experimental results are compared with many related face recognition algorithms, including classic algorithms and the latest algorithms, namely Naive collaborative representation, Naive PALM, Naive LILS, Naive FISTA, PALM, Improved collaborative representation

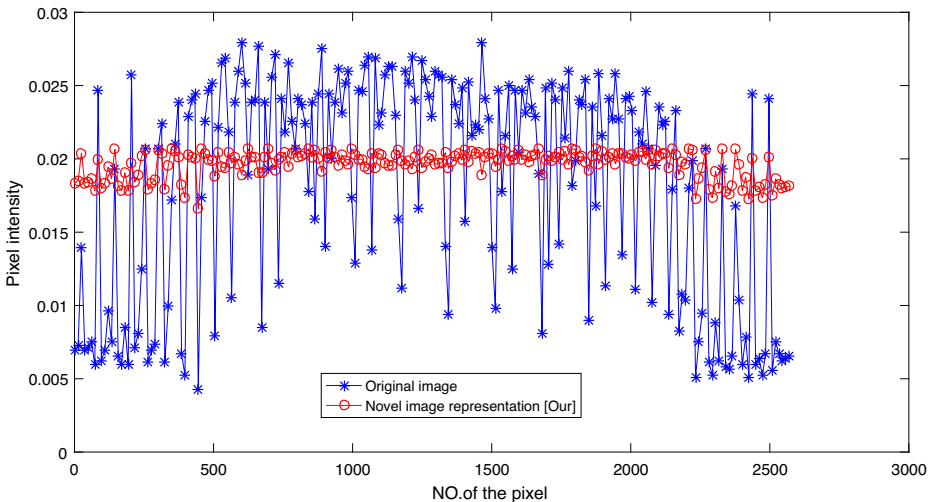


Fig. 5 The standardized data of the first face image of the first subject in the ORL database under the original image and the novel image representation

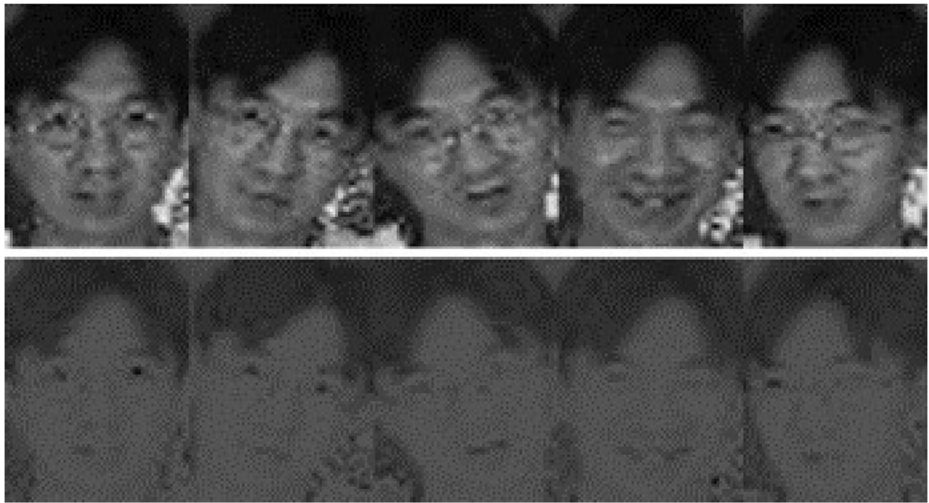


Fig. 6 The original gray face images of a subject in the Georgia Tech face database (the first row) and their corresponding virtual face images (the second row)

[30], BDLRR [37], RSLDA [23], Multi-resolution dictionary learning (MRDL) [15] and Improved image representation and sparse representation (IIRSR) [39]. The algorithm directly applied to the original training samples is called Naive. This section compares the classification error rates of different face recognition algorithms. Experimental results show that the proposed algorithm has achieved satisfactory results. The novel virtual image representation method and weight fusion scheme play an important role in reducing the classification error rate of the algorithm.

In all experiments, the face images of each subject in the face database are divided into training sets and test sets. When fixing the number of training samples for each subject, the program is executed ten times, and the final result is the average of all experimental results. To verify the effect of extended data on the algorithm, we use exponential expansion [19] instead of (11) to generate virtual images, and the experimental results are denoted as Q_{in} [19]. Furthermore, while all algorithms use Euclidean distance by default to calculate the difference between images, we also verified the effect of other distance calculation methods, such as Manhattan distance.

5.1 Experiments on the ORL database

The ORL face database collects face images of subjects from different angles under different lighting conditions. Each subject contains samples of different facial expressions, such as eyes open or closed, smiling or not smiling. At the same time, the face images in the database contain partial occlusions, such as with and without glasses. The face database collected a total of 400 face images of 40 subjects, each subject contains 10 samples, and the size of each face image is 56x46 pixels. Figure 7 shows the face images of three subjects in the face database. On the ORL face database, Table 1 shows the classification error rates of different face recognition algorithms. In the experiment, the number of training samples of each subject is set to 2, 3, 4, and 5 respectively, and the remaining samples are used as the test set.



Fig. 7 Face images of ORL face database

From Table 1, we can clearly see that the proposed algorithm is very competitive in terms of classification accuracy. For example, when the number of training samples is 5, the classification error rate of the proposed algorithm is 2.5% lower than that of “Improved collaborative representation” and 1.5% lower than that of IIRSR2. Similarly, under other training sample numbers, the proposed algorithm has a higher recognition accuracy rate than other face recognition algorithms.

5.2 Experiments on the Georgia Tech database

The Georgia Tech face database has 750 color face images, composed of 50 subjects, and 15 face images for each subject. The face images of each subject are collected under different angles and different illuminations, and there is a small amount of occlusion, such as with or without glasses. We use cropped face images for experiments. The size of the cropped

Table 1 Rate of classification errors (%) on the ORL dataset

Number of training samples per class	2	3	4	5
Our proposed algorithm	11.56	10.00	6.67	6.00
Our proposed algorithm-manhattan	11.25	10.00	6.67	7.00
IIRSR1 [39]	11.25	10.71	7.08	7.00
IIRSR2 [39]	10.94	11.07	6.67	7.50
Improved collaborative representation [30]	12.50	11.43	8.75	8.50
Qin [19]	12.50	11.07	9.58	7.00
Naive collaborative representation	16.56	13.93	10.83	11.50
Naive PALM	19.69	18.57	13.57	13.00
Naive L1LS	19.69	18.93	14.58	13.50
Naive FISTA	18.44	16.79	12.08	13.50
PALM	20.44	11.93	9.88	8.20
MRDL [15]	23.15	18.21	16.62	9.55
RSLDA [23]	24.37	12.85	10.83	8.00
BDLRR [37]	21.87	13.93	10.42	8.00



Fig. 8 Face images of Georgia Tech face database

image is 40x30 pixels and the messy background of the original image is removed. Figure 8 shows the face images of three subjects in the database.

In the experiment, for each subject, the number of training samples is 1, 2, and 3, and the remaining face images are used as the test set. On the Georgia Tech face database, the comparison results of the classification error rates of different face recognition algorithms are shown in Table 2.

From Table 2, compared with other face recognition algorithms, the proposed algorithm obtains the lowest image classification error rate. When the number of training samples for each subject is 3, the proposed algorithm reduces the classification error rate by 4.84% and 1.5% respectively than the “Improved collaborative representation” and the IIRSR2

Table 2 Rate of classification errors (%) on the Georgia Tech dataset

Number of training samples per class	1	2	3
Our proposed algorithm	62.29	50.31	47.33
Our proposed algorithm-manhattan	62.71	50.62	47.17
IIRSR1 [39]	63.00	51.23	48.17
IIRSR2 [39]	63.29	51.69	48.83
Improved collaborative representation [30]	63.86	52.15	52.17
Qin [19]	66.14	56.15	52.83
Naive collaborative representation	66.57	57.54	54.67
Naive PALM	68.14	61.38	58.33
Naive L1LS	68.00	61.08	58.50
Naive FISTA	65.14	54.92	53.50
PALM	65.26	55.71	52.29
MRDL [15]	\	65.86	61.25
RSLDA [23]	70.00	51.70	50.66
BDLRR [37]	67.43	56.77	49.17



Fig. 9 Face images of FERET face database

algorithm. Compared with other face recognition algorithms, the proposed algorithm has greater advantages in classification accuracy.

5.3 Experiments on the FERET database

The face images of the FERET face database are collected from the subjects under different lighting conditions, reflecting the different postures and facial features of the subjects. The experiment uses multiple subsets of the database to verify the face recognition algorithm, which are “ba”, “bj”, “bk”, “be”, “bf”, “bd” and “bg” subsets. In the experiment, there are 1400 face images in all subsets, including 200 subjects, and each subject has 7 face images. The size of each face image in the experiment is 40x40 pixels. Figure 9 shows examples of face images of three subjects in the database. On the FERET database, when the number of training samples for each subject is set to 1, 2, and 3, the classification error rates of different face recognition algorithms are compared to Table 3.

Table 3 Rate of classification errors (%) on the FERET dataset

Number of training samples per class	1	2	5
Our proposed algorithm	52.33	39.30	27.00
Our proposed algorithm-manhattan	52.50	39.50	27.00
IIRSR1 [39]	52.67	39.80	27.25
IIRSR2 [39]	52.08	40.10	28.75
Improved collaborative representation [30]	53.08	39.80	30.25
Qin [19]	56.50	44.70	28.25
Naive collaborative representation	55.67	41.60	31.50
LILS	73.71	58.88	32.93
FISTA	73.48	59.29	33.33
PALM	77.85	69.60	70.25
MRDL [15]	\	56.12	49.02
RSLDA [23]	83.16	55.30	30.25
BDLRR [37]	71.42	54.20	29.75

From Table 3, when the number of training samples for each subject is 1, 2, and 3, the proposed algorithm obtains the lowest classification error rate. When the number of training samples is 3, the proposed algorithm reduces the classification error rate by 3.25% and 1.75%, respectively, compared with the “Improved collaborative representation” algorithm and the IIRS2 algorithm. Experimental results show that the proposed algorithm has better image classification results.

From the experiments, compared with the other two databases, the recognition accuracy of the algorithm is higher on the ORL database. The main reason is that the face images in the Georgia Tech database and the FERET database vary greatly under different expressions, lighting, and postures, and the differences between different face images of the same subject are greater. Also, when calculating the distance between images, using Manhattan distance and Euclidean distance achieved similar results.

6 Conclusions

In this paper, we propose an improved image representation method and design a weight fusion scheme, based on both of which implement a novel image classification algorithm and apply it to face recognition. This image representation method converts the original image into a virtual image, effectively preserving the large-scale information and global features of the original image. In addition, the virtual image can serve as a complementary representation of the object. Combining the original image can represent objects from different perspectives, which plays an important role in the correct classification of the algorithm, especially deformable objects such as faces. Moreover, the proposed weight fusion scheme is simple and effective, which is used to fuse the original image and the virtual image and obtain the final classification distance of the test image. The experimental results show the rationality and feasibility of the proposed algorithm. Compared with other related face recognition algorithms, the proposed algorithm achieves the best classification accuracy, and the algorithm can also be applied to other image classification tasks.

References

1. Aharon M, Elad M, Bruckstein A (2006) K-SVD : an algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans Signal Process* 54:4311–4322
2. Deng W, Hu J, Guo J (2012) S.R.C. extended, undersampled face recognition via intraclass variant dictionary. *IEEE Trans Pattern Anal Mach Intell* 34(9):1864–1870
3. Deng W, Hu J, Guo J (2018) Face recognition via collaborative representation: its discriminant nature and superposed representation. *IEEE Trans Pattern Anal Mach Intell* 40(10):2513–2521
4. Ding C, Xu C, Tao D (2015) Multi-task pose-invariant face recognition. *IEEE Trans Image Process* 24(3):980–993
5. Fan Z, Wei C (2020) Fast kernel sparse representation based classification for undersampling problem in face recognition. *Multimed Tools Appl* 79:7319–7337
6. Fang XZ, Xu Y et al (2015) Learning a nonnegative sparse graph for linear regression. *IEEE Trans Image Process* 24(9):2760–2771
7. Gao Y, Ma J, Yuille AL (2017) Semi-supervised sparse representation based classification for face recognition with insufficient labeled samples. *IEEE Trans Image Process* 26(5):2545–2560
8. Jiang Z, Lin Z, Davis LS (2013) Label consistent k-svd : learning a discriminative dictionary for recognition. *IEEE Trans Pattern Anal Mach Intell* 35(11):2651–2664
9. Kafai M, An L, Bhanu B (2014) Reference face graph for face recognition. *IEEE Press*, Hoboken
10. Keinert F, Lazzaro D, Morigi S (2019) A robust group-sparse representation variational method with applications to face recognition. *IEEE Trans Image Process* PP(6):1–1

11. Li Z, Lai Z, Xu Y et al (2017) A locality constrained and label embedding dictionary learning algorithm for image classification. *IEEE Trans Neural Netw Learn Syst* 28(2):278–293
12. Liao M, Gu X (2020) Face recognition approach by subspace extended sparse representation and discriminative feature learning. *Neurocomputing* 373:35–49
13. Lin G, Yang M, Shen L et al (2018) Robust and discriminative dictionary learning for face recognition. *Int J Wavelets Multiresolut Inf Process* 16:1840004
14. Lin G, Yang M, Yang J, Shen L, Xie W (2018) Robust, discriminative and comprehensive dictionary learning for face recognition. *Pattern Recognit* 81:341–356
15. Luo X, Xu Y, Yang J (2019) Multi-resolution dictionary learning for face recognition. *Pattern Recogn* 93:283–292
16. Ma X, Zhang F, Li Y et al (2018) Robust sparse representation based face recognition in an adaptive weighted spatial pyramid structure. *Sci China Inf Sci* 61:012101
17. Moeini A, Moeini H (2015) Real-world and rapid face recognition toward pose and expression variations via feature library matrix. *IEEE Trans Inf Forensics Secur* 10(5):969–984
18. Pengyue Z, Xinge Y, Weihua O et al (2016) Sparse discriminative multi-manifold embedding for one-sample face identification. *Pattern Recognit* 52:249–259
19. Qian R (2018) Inverse transformation based weighted fusion for face recognition. *Multimed Tools Appl* 77:28441–28456
20. Samaria FS, Harter AC (1994) Parameterisation of a stochastic model for human face identification. In: *Proceedings of 1994 IEEE workshop on applications of computer vision*. IEEE Comput. Soc. Press, pp 138–142
21. Shrivastava A, Patel VM, Chellappa R (2015) Non-linear dictionary learning with partially labeled data. *Pattern Recogn* 48(11):3283–3292
22. Tan J, Zhang T, Zhao L, Luo X et al (2021) A robust image representation method against illumination and occlusion variations. *Image Vis Comput* 112(4):104212
23. Wen J, Fang X, Cui J et al (2019) Robust sparse linear discriminant analysis. *IEEE Trans Circuits Syst Video Technol* 29(2):390–403
24. Wright J, Yang AY, Ganesh A, Sastry SS, Ma Y (2009) Robust face recognition via sparse representation. *IEEE Trans Pattern Anal Mach Intell* 31(2):210–227
25. Wu M, Wang S, Li Z et al (2021) Joint latent low-rank and non-negative induced sparse representation for face recognition. *Appl Intell* 51:8349–8364
26. Xu Y, Zhang D, Yang J et al (2011) A two-phase test sample sparse representation method for use with face recognition. *IEEE Trans Circuits Syst Video Technol* 21(9):1255–1262
27. Xu Y, Zhu X, Li Z et al (2013) Using the original and 'symmetrical face' training samples to perform representation based two-step face recognition. *Pattern Recogn* 46(4):1151–1158
28. Xu Y, Li X, Yang J, Zhang D (2014) Integrate the original face image and its mirror image for face recognition. *Neurocomputing* 131:191–199
29. Xu Y, Li XL, Yang J et al (2014) Integrating conventional and inverse representation for face recognition. *IEEE Trans Cybern* 44(10):1738–1746
30. Xu Y, Zhang B, Zhong Z (2015) Multiple representations and sparse representation for image classification. *Pattern Recogn Lett* 68:9–14
31. Xu Y, Zhong Z, Yang J, You J, Zhang D (2017) A new discriminative sparse representation method for robust face recognition via $l(2)$ regularization. *IEEE Trans Neural Netw Learn Syst* 28(10):2233–2242
32. Xu Y, Li Z, Tian C, Yang J (2019) Multiple vector representations of images and robust dictionary learning. *Pattern Recogn Lett* 128:131–136
33. Yang J, Luo L, Qian J, Tai Y, Zhang F, Xu Y (2017) Nuclear norm based matrix regression with applications to face recognition with occlusion and illumination changes. *IEEE Trans Pattern Anal Mach Intell* 39(1):156–171
34. Ye MJ, Hu CH, Wan LG et al (2021) Fast single sample face recognition based on sparse representation classification. *Multimed Tools Appl* 80:3251–3273
35. Zhang Q, Li B (2010) Discriminative k-svd for dictionary learning in face recognition. In: *2010 IEEE Computer society conference on computer vision and pattern recognition*. IEEE, pp 2691–2698
36. Zhang L, Yang M, Feng X (2011) Sparse representation or collaborative representation: which helps face recognition? In: *Proceedings of IEEE international conference on computer vision*, pp 471–478
37. Zhang Z, Xu Y, Shao L, Yang J (2018) Discriminative blockdiagonal representational learning for image recognition. *IEEE Trans Neural Netw Learn Syst* 29(7):3111–3125

38. Zhang Y, Zheng S, Zhang X et al (2020) Multi-resolution dictionary learning method based on sample expansion and its application in face recognition. *SIVip*
39. Zheng S, Zhang Y, Liu W et al (2020) Improved image representation and sparse representation for image classification. *Appl Intell* 50:1687–1698

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.