TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 08/12 pp62-67
Volume 18, Number 1, February 2013

Sparse Representation for Face Recognition Based on Constraint Sampling and Face Alignment

Jing Wang, Guangda Su*, Ying Xiong, Jiansheng Chen, Yan Shang, Jiongxin Liu, and Xiaolong Ren

Abstract: Sparse Representation based Classification (SRC) has emerged as a new paradigm for solving recognition problems. This paper presents a constraint sampling feature extraction method that improves the SRC recognition rate. The method combines texture and shape features to significantly improve the recognition rate. Tests show that the combined constraint sampling and facial alignment achieves very high recognition accuracy on both the AR face database (99.52%) and the CAS-PEAL face database (99.54%).

Key words: classification; face recognition; feature extraction; face alignment

1 Introduction

Face recognition has been one of the most challenging research areas in the world. Many face recognition methods have been proposed, such as Principal Components Analysis (PCA)[1], Independent Components Analysis (ICA)[2], the Hausdorff distance[3], Elastic Graph-Matching (EGM)[4], and Support Vector Machines (SVM)[5]. These methods achieve satisfactory results in well-controlled environments, but their accuracy seriously degrades in uncontrolled or moderately controlled environments, especially when the facial pose and illumination of the input image deviate too much from those of the training images[6]. Recently, the Sparse Representation based Classification (SRC) has attracted much attention due to its effectiveness for face recognition with significant illumination and expression variations[7]. This method uses the theory of compressive sampling[8] to exploit the discriminative nature of the sparse representation for classification. With proper selection of the training samples and the number of features, the SRC algorithm achieves good recognition results even with serious variations in the illumination conditions or expressions. The SRC algorithm has given promising recognition results on public face databases including the extended Yale face database B[9] and the AR face database[10].

Wright et al.[7] claimed that if the recognition sparsity was properly used, the choice of features would no longer be critical. They then proposed a less-structured feature called Randomfaces and validated through experiments that this could achieve similar results as conventional features. However, different parts of a face contain different amounts of information. If the features are selected for high-informational areas, the recognition efficiency will be improved for the same feature dimension. Thus, this paper presents a feature extraction method called *constraint sampling* that obtains a *fixedface* feature value using key facial points. Comparison with results using traditional features demonstrates that a proper uneven selection of features will capture the main characteristics and achieve better recognition performance. Another advantage of constraint sampling is that the face images are better aligned through the key point locating process, so it efficiently overcomes the alignment

• Jing Wang, Guangda Su, Jiansheng Chen, Yan Shang, and Xiaolong Ren are with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China. E-mail: jwang08@mails.tsinghua.edu.cn; susu@tsinghua.edu.cn.
• Ying Xiong is with the Department of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA. E-mail: yxiong@seas.harvard.edu.
• Jiongxin Liu is with the Department of Computer Science, Columbia University, New York, NY 10027, USA. E-mail: liujiongxin@gmail.com.
∗ To whom correspondence should be addressed.
Manuscript received: 2012-07-16; accepted: 2012-12-05

limitation[11] of SRC.

This study also demonstrates that the face recognition can be significantly improved by using shape information. The coordinates of the facial feature points are concatenated as a shape feature to be used by the SRC algorithm. Biometric fusion is then used to combine the texture feature and the shape feature recognition results. Tests show that including the shape feature significantly improves the recognition performance.

2 Sparse Recognition Classification Method

The SRC classifier was proposed by Wright et al.[7] The fundamental assumption of this method is that the training samples from a single class lie on a linear subspace. Therefore, given sufficient training samples for the i-th object class $\boldsymbol{A}_i = [v_{i,1}, v_{i,2}, \cdots, v_{i,n_i}] \in \mathbf{R}^{m \times n_i}$, any new test sample, $\boldsymbol{y} \in \mathbf{R}^m$, from the same class lies within the linear span of the training samples associated with object i:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}_0 + z \tag{1}$$

where $\boldsymbol{x}_0 = [0, \cdots, 0, \alpha_{i,1}, \alpha_{i,2}, \cdots, \alpha_{i,n_i}, 0, \cdots, 0]^{\mathrm{T}} \in \mathbf{R}^n$ is an unknown coefficient vector whose entries should be zero except for those associated with the i-th class, $\boldsymbol{A}$ is defined for the entire training set as the concatenation of the n training samples of all k object classes, and $z \in \mathbf{R}^m$ is a noise term with bounded energy $\|z\|_2 < \varepsilon$. The sparse solution, $\boldsymbol{x}_0$, can be approximately recovered by solving the following stable l^1-minimization problem.

$$(l^1) : \widehat{x} = \arg\min \|\boldsymbol{x}\|_1, \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2 \leqslant \varepsilon \tag{2}$$

This problem can be solved in polynomial time by standard convex programming methods[12].

SRC uses all the samples in the training set during each recognition time to obtain better results than the Nearest Neighbor (NN) and Nearest Sub-space (NS) algorithms, which only use part of the training set for recognition[13]. Since the feature choice is no longer critical, Wright et al.[7] claimed that PCA[1], ICA[2], DownSampled face[7], and even RadonFace[7] can be chosen as features. Nevertheless, since sparsity can not be satisfied in practical applications, the feature choice again becomes critical and careful feature selection will greatly improve the results.

3 Face Recognition Methods

Two methods that integrate SRC are introduced in this section. The first method uses *constraint sampling* with the *fixedface* attribute to overcome the limitations of the SRC method for satisfactory recognition results. The second method combines Constraint Sampling and Facial Alignment (CSFA-SRC) to achieve the best performance among all the presented methods.

3.1 SRC with Constraint Sampling (CS-SRC)

Wright et al.[7] showed that proper use of the sparsity for recognition problem reduces the importance of the feature choice. However, the feature choice will always be important since most facial information is located around the eyes and the mouth[14-16]. The importance of feature selection is shown here using the constraint sampling feature extraction method. The key facial points are used to select fixed features in specific areas. Constraint sampling pays more attention to facial areas around the eyes, nose, and mouth that contain more information for identification and neglects areas with small gradient change that are sensitive to Gaussian noise. This non-uniform unequal selection strategy then identifies the main characteristics to give better recognition performance. The main limitation of the SRC algorithm is the need for accurate pixel-level alignment between the test image and the training set. This leads to vulnerability to misalignment, making it inappropriate for deployment outside a laboratory setting[11]. Constraint sampling combines alignment and feature extraction into one procedure to effectively overcome this limitation.

The fixedface is defined from features generated through constraint sampling, including the eye, nose, and mouth areas generated on the basis of key points. The eyeballs, nose tip, and mandibular points shown in Fig. 1 are used as the key points.

The eye, nose, mouth areas are then cropped according to these four key points. These areas are

Fig. 1 Key facial points.

then evenly downsampled according to a predetermined downsampling ratio, h. The fixedface feature is obtained by concatenating the downsampled images into a vector.

Thus, fixedface is well aligned and more selective than traditional features. This constraint sampling can then be used with SRC to achieve satisfactory recognition results, as will be shown in Section 4.

3.2 Sparse representation for face recognition based on CSFA-SRC

Both ASM[17,18] and AAM[12] are Point Distribution Models (PDM), which exploit a linear formulation of PDMs in an iterative search procedure to locate the modeled structures in the test image. The successful application of ASM and AAM for facial alignment suggests that the shape points of human faces align on a linear space. Thus, SRC is extended here to process shape information extracted from facial contours.

105 feature points are chosen to describe the facial shape as shown in Fig. 2. The points are ranked and concatenated according to their x and y coordinates to give a 210-dimensional vector: $(x_0,\ y_0,\ x_1,\ y_1,\cdots,\ x_{104},\ y_{104})$, where $(x_i,\ y_i)$ are the coordinates of the i-th feature point. Rotation, zoom, and translation are used to normalize the vector so that the y-coordinates of the two eyeballs are the same, the x-coordinate difference between the two eyeballs is 10 pixels, and the nose tip is set as the origin. The ASM+AAM model[19] is used to locate the feature points.

The fixedface result is used as the texture information. The shape information is computed as the coordinates of the facial feature points. The texture and shape features are then used by the modified SRC algorithm and fused as a weighted sum. For the fusion algorithm, the modified SRC algorithm analyzes the texture feature to produce the similarities R_{1i} (i=1, $\cdots$, k, for k classes) (defined in Eq. (5) in Algorithm

Fig. 2 105 feature points on the frontal face.

1) as a result and analyzes the shape feature to produce the similarities R_{2i} (i=1, $\cdots$, k) as an equivalent result. A similarity measurement symbol S_i (i=1, $\cdots$, k) is then defined as

$$s_i = b_1 \times R_{1i} + b_2 \times R_{2i} \tag{3}$$

where $b_1+b_2 = 1, 0 \leqslant b_1 \leqslant 1$. The test image identity, $\boldsymbol{I}$, is calculated as

$$\text{identity}(\boldsymbol{I}) = \arg\max_i s_i \tag{4}$$

The modified SRC procedure given in Algorithm 1 has a different output from the original SRC algorithm with the output of the similarity of the test sample to each training class instead of the test image identity.

Algorithm 1 Modified SRC Procedure

(1) Inputs: a matrix of training samples $\boldsymbol{A} = [\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_k] \in \mathbf{R}^{m\times n}$ for k classes, a test sample $\boldsymbol{y} \in \mathbf{R}^m$, and an optional error tolerance $\varepsilon > 0$.

(2) Normalize the columns of $\boldsymbol{A}$ to have a unit l^2-norm.

(3) Solve the l^1-minimization as in Eq. (2).

(4) Compute the residuals $r_i(\boldsymbol{y}) = \left\|\boldsymbol{y} - \boldsymbol{A}\delta_i(\widehat{\boldsymbol{x}})\right\|_2$ for i=1, $\cdots$, k. (For any $\boldsymbol{x} \in \mathbf{R}^n$, $\delta_i(\boldsymbol{x}) \in \mathbf{R}^n$ is a new vector whose only nonzero entries are the entries in $\boldsymbol{x}$ associated with class i.)

(5) Output: the similarity from each class:

$$R_i(\boldsymbol{y}) = \min_{1\leqslant j\leqslant k} \frac{r_j(\boldsymbol{y})}{r_i(\boldsymbol{y})} \tag{5}$$

4 Test Results

Two sets of tests were performed. The first set used a series of interactive recognition tests on public databases. The feature points were calibrated manually to verify the effectiveness of the constraint sampling and shape features. The second set of tests used the automatic feature recognition without manual adjustments. The whole process was done automatically to validate the usefulness of this method for practical face recognition applications.

4.1 Interactive recognition experiments

4.1.1 Face recognition results on the AR face database

The AR database consists of over 4000 frontal images of 126 individuals. 26 pictures were taken for each individual in two separate sessions. These images include facial variations including illumination changes, expression changes, and facial disguises. The tests use a subset of the dataset consisting of 45 male subjects and 45 female subjects. The tests used 14 images for each subject with only illumination and expression changes. The 14 selected images included seven Session 1 images for training and seven Session 2 images for testing. Five feature space dimensions of 48, 192, 432, 768, and 1200 were selected. The images were all converted to grayscale.

Since the performance of the traditional SRC algorithm is not dependent on traditional features[1], the SRC algorithm was tested using downsampled images as a baseline. Then, the SRC was used with constraint sampling (CS-SRC) with four key points that were marked manually. Then, the CSFA-SRC algorithm was used with the feature points marked manually to show its theoretical performance. The fusion rate, b_i, in Eq. (3) was 0.6:0.4. (The recognition performance was tested for different fusion rates with the highest performance achieved for rates between 0.5-0.7.) The test results are shown in Fig. 3.

As shown in Fig. 3, the traditional SRC (TSRC) achieves a recognition rate of 93.49% with 432- and 768- dimensional Randomfaces and the recognition rate does not always increase as the dimension grows. For feature dimensions from 48 to 432, the recognition rate increases as the dimension increases; however, the recognition rate drops slightly for dimensions larger than 763. Since only increasing the dimension of the traditional texture features does not provide more useful information, but actually makes the function more sensitive to registration, which reduces the recognition rates. Other types of features such as the shape feature should be used to improve the recognition rate. Figure 3 shows that the recognition rate is significantly improved by including the shape information with CSFA-SRC achieving a recognition rate of 99.52% with the 192-dimensional fixedface. Only three of the 630 test facial images were wrongly classified. The excellent recognition rate of this algorithm demonstrates the importance of proper feature selection.

4.1.2 Face recognition results on the CAS-PEAL face database

The CAS-PEAL face database[20] contains 99 594 images of 1040 individuals with varying Poses, Expressions, Accessories, and Lighting (PEAL). The expression library contains 377 individuals with each individual have 6 different expressions. The expression library is used here to evaluate the algorithm recognition rate. For each individual, 4 of the facial images were randomly selected for training with the other 2 used for testing. This database is substantially more challenging than the AR facial database with 202 subjects and only 4 training images per subject. The five feature space dimensions of 48, 192, 432, 768, and 1200 were used for the tests. The tests were similar to those used on the AR face database with the results shown in Fig. 4.

The results in Fig. 4 show that the traditional SRC achieves a maximum recognition rate of 93.56% with a feature space dimension of 192. As before, the recognition rates do not always increase as the dimension increases but decreases as the dimension increases from 192 to 1200. The CS-SRC results are much better than those of the traditional SRC, with a maximum recognition rate of 98.02% for a feature dimension of 192. CSFA-SRC gives the best rate with a maximum recognition rate of 99.54% with only two test samples among the 404 samples wrongly classified.

Thus, these tests on both the AR face database and the CAS-PEAL face database show that constraint sampling is an effective feature selection method that improves the recognition rate of the SRC algorithm and that the face recognition rate can be significantly

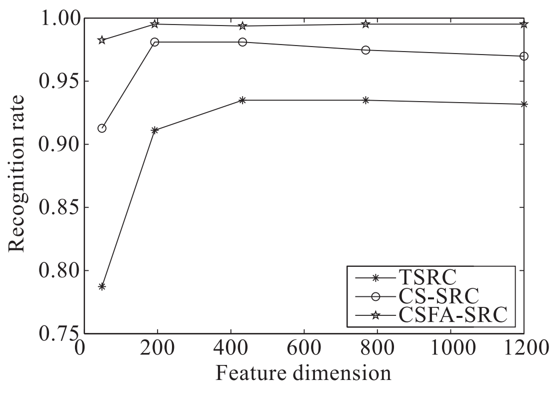


Fig. 3 Recognition rates on the AR face database.

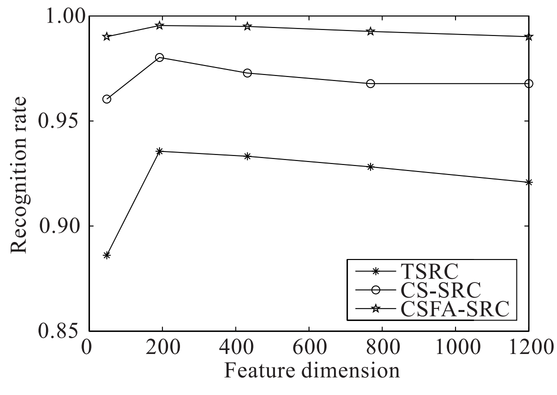


Fig. 4 Recognition rates on the CAS-PEAL face database.

improved by using shape information.

4.2 Automatic recognition experiments

The effectiveness of the entire system is illustrated by applying the CSFA-SRC algorithm to the AR and the CAS-PEAL face databases with each step running automatically. The recognition times with CSFA-SRC are almost the same as with SRC. The recognition results are shown in Fig. 5.

The automatic CSFA-SRC recognition rate is not as good as with manual intervention, since the feature locating algorithm is not robust enough with illumination and expression changes. However, the automatic algorithm significantly outperforms the traditional SRC with maximum recognition rates of 98.25% on the AR face database and 98.02% on the CAS-PEAL face database.

5 Conclusions

This paper presents two improved feature selection methods. The algorithms significantly improve the limitations of the SRC algorithm by combining texture and geometric information for the face recognition. Constraint sampling based on the key points is used to better align the images input to the SRC algorithm. Then, the coordinates of the facial feature points are chosen for the shape feature, which significantly improve the SRC recognition rate. Different parts of the face contain different amount of information, so the constraint sampling with the combined texture and shape characteristics give more information and achieve better results. The excellent recognition rates of the algorithms demonstrate the importance of feature selection for SRC algorithms.

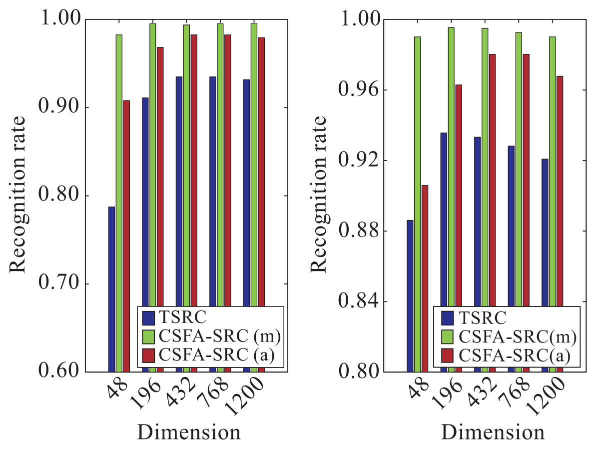


Fig. 5 Automatic face recognition: a) recognition rates on the AR face database; b) recognition rates on the CAS-PEAL face database. CSFA-SRC(m) refers to CSFA-SRC with manually marked feature points, while CSFA-SRC(a) refers to the totally automated CSFA-SRC.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Nos. 60772047 and 61101152), the National Science & Technology Pillar Program during the Eleventh Five-year Plan Period (No.2006BAK08B07), and the Chuanxin Foundation from Tsinghua University (No. 110107001).

References

[1] M . Turk and A. Pentland, Eigenfaces for recognition, *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp.71-86, 1991.

[2] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, Face recognition by independent component analysis, *IEEE Trans. Neural Networks*, vol. 13, no. 6, pp. 1450-1464, 2002.

[3] B. Takacs, Comparing face images using the modified Hausdorff distance, *Pattern Recognition*, vol. 31, no. 12, pp. 1873-1881, 1998.

[4] L. Wiskott, J. M. Fellous, N. Kruger, and C. V. D. Malsburg, Face recognition by elastic bunch graph matching, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 775-779, 1997.

[5] O. Deniz, M. Castrillon, and M. Hernandez, Face recognition using independent component analysis and support vector machines, *Pattern Recognition Letters*, vol. 24, no. 13, pp. 2153-2157, 2003.

[6] P. N. Belhumeur and D. J. Kriegman, What is the set of images of an object under all possible illumination conditions, *Int. Journal of Computer Vision*, vol. 28, no. 3, pp. 245-260, 1998.

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

[8] E. J. Candes and M. B. Wakin, An introduction to compressive sampling, *IEEE Signal Progressing Magazine*, vol. 25, no. 2, pp. 21-30, 2008.

[9] K. C. Lee, J. Ho, and D. Kriegman, Acquiring linear subspaces for face recognition under variable lighting, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27, no. 5, pp. 684-698, 2005.

[10] A. Martinez and R. Benavente, The AR face database, *CVC Tech. Report*, vol. 24, 1998.

[11] A. Wagner, J. Wright, A. Ganesh, Z. Zhou, and Y. Ma, Towards a practical face recognition system: Robust registration and illumination by sparse representation, presented at the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 2009.

[12] T. F. Cootes, G. J. Edwards, and C. J. Taylor, Active appearance models, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681-685, 2001.

[13] Y. Li, B. Chen, X. Zhang, and H. Shu, Face recognition algorithm by using sparse representation and Tchebichef moments, (in Chinese), *Journal of Southeast University*, vol. 42, no. 2, pp. 249-253, 2012.

[14] F. Nahm, A. Perret, D. Amaral, and T. Albright, How do monkeys look at faces, *Journal of Cognitive Neuroscience*, vol. 9, no. 5, pp. 611-623, 1997.

[15] I. Biederman, Human image understanding: Recent research and theory, *Computer Vision Graphics and Image Processing*, vol. 32, no. 1, pp. 29-73, 1985.

[16] W. E. L. Grimson and J. R. Meyer-Arendt, *Object Recognition by Computer—The Role of Geometric Constraints*, MIT Press, 1990.

[17] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, Active shape models-their training and application, *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38-59, 1995.

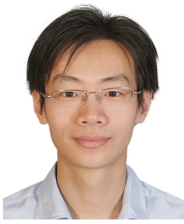
[18] T. F. Cootes, C. J. Taylor, and A. Lanitis, Multi-resolution search with active shape models, presented at the 12th IAPR International Conference on Pattern Recognition, Jerusalem, Israel, 1994.

[19] J. Wang, G. Su, J. Liu, and X. Ren, Facial feature point location algorithm combining improved ASM and AAM, (in Chinese), *Journal of Optoelectronics Laser*, vol. 22, no. 8, pp. 1227-1230, 2011.

[20] G. Wen, The CAS-PEAL large-scale Chinese face database and baseline evaluations, *IEEE Trans. System Man*, vol. 38, no. 1, pp. 149-161, 2008.

Jing Wang is a PhD candidate in the Department of Electronic Engineering of Tsinghua University. She received her BS degree from Nanjing University in 2008. She has published 5 papers, and her research interests including face recognition, feature extraction, and face alignment.

Guangda Su is a professor in Department of Electronic Engineering of Tsinghua University. He received his BEng degree from Tsinghua University in 1977. His research interest is face recognition and image processing.



Ying Xiong received his BEng degree from Tsinghua University and MEng degree from Harvard University in 2010 and 2012. He is currently a PhD candidate at Harvard University. He has worked as research assistant at NEC-Labs America and software engineering intern at Google Inc. His major research interest is computer vision.

Jiansheng Chen received his BEng and MEng degree from Tsinghua University in 1999 and 2002, and received his PhD degree from the Chinese University of Hong Kong in 2006. He is currently an assistant professor in Tsinghua University. His major research interest is pattern recognition and image processing.

Yan Shang received her BEng and MEng degree from Tianjing University in 1993 and 1996. Now she works at Tsinghua University. Her major research interest is pattern recognition and image processing.

Jiongxin Liu received his BEng and MEng degree from Tsinghua University in 2007 and 2009. He is currently a research assistant in Computer Science Department (pursuing PhD degree) at Columbia University. His major research interest is image processing and computer vision.

Xiaolong Ren received his BEng and MEng degree from Tsinghua University in 2005 and 2009. He is currently an engineer in Tsinghua University. His major research interest is face recognition and image processing.