Computer Aided Drafting, Design and Manufacturing
Volume 25, Number 2, June 2015, Page 14

CADDM

Face Recognition Algorithm Fusing Monogenic Binary Coding and Collaborative Representation

FU Yu-xian, PENG Liang-yu

College of Physics and Information Science , Hunan Normal University, Changsha, 410081, China.

Abstract: Monogenic binary coding (MBC) have been known to be effective for local feature extraction, while sparse or collaborative representation based classification (CRC) has shown interesting results in robust face recognition. In this paper, a novel face recognition algorithm of fusing MBC and CRC named M-CRC is proposed; in which the dimensionality problem is resolved by projection matrix. The proposed algorithm is evaluated on benchmark face databases, including AR, PolyU-NIR and CAS-PEAL. The results indicate a significant increase in the performance when compared with state-of-the-art face recognition methods.

Key words: face recognition; monogenic binary coding; collaborative representation

1 Introduction

Face recognition, as one of the most typical applications of image analysis and understanding, has been extensively studied in the past two decades[1]. However, it is still difficult for machines to recognize human faces accurately under the uncontrolled circumstances, including occlusions and variations in pose, illumination, expression and aging, etc. Various methods have been proposed for face feature extraction, such as Eigenfaces[2], Fisherfaces[3], Gabor Featutre based Classification[4], LBP[5], LGBP[6] and MBC[7], etc. However, the performance of Eigenfaces and Fisherfaces degrades when the lighting conditions or expression change. Literature[4] demonstrated the promising performance of Gabor feature, but with significant time and space complexity. LBP methods use local structural information and histogram of sub-regions to extract facial features. In addition to its efficiency, the simplicity of LBP allows for faster feature extraction than other methods. The advantage of MBC over the Gabor is that it has much lower time and space complexity but with better performance. The advantage of MBC over the LBP is that it has competitive performance. This is mainly because monogenic signal analysis is a compact representation of features with little information loss. The recognition of a query face image is usually accomplished by classifying the features extracted from this image. The most popular classifier for FR may be the nearest neighbor (NN) classifier due to its simplicity and efficiency. SVM[8] can be also seen a classifier.

Wright et al.[9] reported a very interesting work by using sparse representation for robust face recognition (FR). The testing image is coded as a sparse linear combination of the training samples, and the representation fidelity is measured by the l_1-norm or l_2-norm of the coding residual. Such a sparse representation based classification (SRC) scheme achieves a great success in FR, and it boosts the research of sparsity based pattern classification. But, the sparsity constraint on the coding coefficients makes SRC's computational cost very high. In [10], Zhang et al. indicated that the success of SRC actually comes from its collaborative representation of query face image over all classes of training samples and proposed a very simple yet much more efficient face classification scheme, namely CR based classification with regularized least square (CRC_RLS). The SRC and CRC_RLS scheme is a very powerful classifier. The method[11-12] aims to combine LBP and sparse representation together, which demonstrated the effectiveness of fused algorithm. Motivated by the success of method[11-12], a new face recognition algorithm fusing monogenic binary coding and collaborative representation is proposed in this paper. Zhang et al.[13] proposed an unsupervised learning method for dimensionality reduction in SRC, and it leads to higher FR rates than PCA and random

Corresponding author: FU Yu-xian, Male, **E-mail:** 363533517@qq.com.

projection. This validates that a well designed dimensionality reduction method can benefit the sparse classification scheme. So the method[13] is used to reduce the dimensionality of histogram features in this paper.

The rest of the paper is organized as follows. Section 2 briefly describes the MBC algorithm. Section 3 presents in detail the CRC_RLS algorithm. Section 4 details the whole scheme of face recognition algorithm fusing monogenic binary coding and collaborative representation. Section 5 presents the experimental results, and Section 6 concludes the paper.

2 Monogenic Binary Coding (MBC)

Monogenic signal was introduced by Felsberg and Sommer in 2001[14] to generalize the analytic signal from 1D to 2D. The monogenic representation of 2D signals is accomplished via the Riesz transform[15]. Monogenic signal representation decomposes a face image signal into three complementary components: amplitude, phase and orientation. The local amplitude a, the local orientation o and the local phase p can be computed by Eq. 1.

$$\begin{cases} a = \sqrt{g^2 + h_x^2 + h_y^2} \\ p = -\operatorname{sign}\left(h_x\right)\arctan\left(\dfrac{\sqrt{h_x^2 + h_y^2}}{g}\right) \\ o = \arctan\left(\dfrac{h_y}{h_x}\right) \end{cases} \tag{1}$$

with

$$\begin{cases} g = I * F^{-1}\left(G\left(w\right)\right) \\ h_i = F^{-1}\left(\dfrac{\sqrt{-1}w_i}{\sqrt{w_x^2 + w_y^2}} \cdot H\right), i \in \left\{x, y\right\} \end{cases}$$

Here, I denotes the input image, and * denotes the convolution operator, $G(w)$ denotes the Log-Gabor filter in Fourier domain, F^{-1} represents the 2D inverse Fourier transform, $H=F(g)$, The frequency response of log-Gabor filters can be described as

$$G\left(w\right) = \exp\left\{-\frac{\left[\ln\left(\dfrac{w}{w_0}\right)\right]^2}{\sqrt{2}\ln\left[\ln\left(\dfrac{\sigma}{w_0}\right)\right]}\right\}$$

where w_0 is the center frequency and σ is the scaling factor of the bandwidth. w_x and w_y are the horizontal and vertical frequencies. Fig. 1 shows a face image. Fig. 2 shows the monogenic representation of a face image at one scale. We can see that the facial local structures are well captured in its monogenic components.

Fig. 1. A face image.

(a) (b) (c)

Fig. 2. Monogenic representation of a face image at one scale: (a) amplitude component, (b) orientation component (c) phase component.

Monogenic binary coding contains two parts. The first part encodes the variation between the central pixel and its surrounding pixels in a local patch, while the second part encodes the value of central pixel itself. Like LBP[5], the local variation of monogenic amplitude could be coded by comparing the amplitude value of central location with those of its neighbors. The information of central pixel itself can have discriminative information which may not be carried out by the local variation. For instance, the two pixels having the same local variation pattern may have very different intensities. Yang[7] proposed to use the imagery part of monogenic signal representation to encode the local feature intensity information of the central pixel. Given the significant time complexity of the proposed algorithm, the monogenic orientation and phase information is not used in this article. Fig. 3 shows the MBC of a face image at three scales.

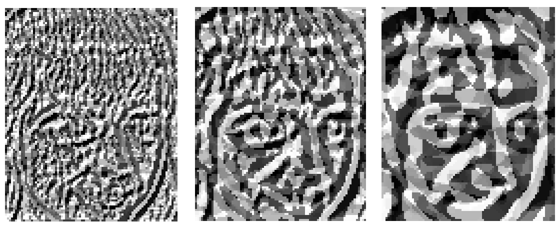

Fig. 3. MBC of a face image at three scales.

The statistical information of image local areas can be described by local histograms, which are robust to the image occlusion and variations of pose, expression, and noise, etc. For each kind of pattern map on each scale, it is partitioned into multiple non-overlapping regions, and then the local histogram is built for each sub-region. Finally, all the local histograms across different scales and different regions are concatenated into a single histogram vector to represent the face image.

3 Collaborative Representation Using Regularized Least Square (CRC-RLS)

CRC-RLS proposed by Zhang et al.[10] is a very simple yet very effective face recognition method that exploits the role of collaboration between classes in representing the query sample. Denote by $A_i \in \Re^{m\times n}$ the dataset of the ith class, and each column of A_i is a sample of class i. Suppose that we have K classes of subjects, and let A=[A_1, A_2,…,A_K]. The query sample is denoted $y \in \Re^m$. In order to collaboratively represent the query sample using A with low computational burden, Zhang[8] proposed to use the regularized least square method.

There is

$$\hat{\rho} = \arg\min_{\rho} \left\{ \left\| y - A\rho \right\|_2^2 + \lambda \left\| \rho \right\|_2^2 \right\} \tag{2}$$

where λ is the regularization parameter. The solution of CR with regularized least square in Eq. 2 can be easily and analytically derived as

$$\hat{\rho} = \left(A^T A + \lambda I \right)^{-1} A^T y \tag{3}$$

Let $P=(A^TA+\lambda I)^{-1}A^T$. Clearly, P is independent of y so that it can be pre-calculated as a projection matrix. Once a query sample y comes, we can just simply project y onto P via Py. This makes CR very fast.

4 Face Recognition Algorithm Fusing Monogenic Binary Coding and Collaborative Representation

It can be seen that when the 8 closest neighbors of a pixel are involved in local variation coding, each MBC pattern will have 10 bits. Then the number of possible patterns for each MBC is 1024, which is larger than that of previous binary coding methods such as LBP, LGBP. It is easy to see that multi-scale monogenic signal representations are redundant. We adopt the SDR (Sparse Dimensionality Reduction)[16] scheme to reduce the histogram feature dimension while enhancing its discrimination. In [16], Denote by $z_i \in R^{m\times 1}$ the i-th training sample of A and by A_i= [z_1,…z_{i-1}, z_{i+1}, …z_n] $\in R^{m\times(n-1)}$ the collection of training samples without the i-th sample, an orthogonal DR matrix P was learnt under the framework of sparse representation, and it achieves better performance than Eigenfaces and Randomfaces in the CRC scheme. Specifically, the matrix P is learnt via the following objective function based on Leave-One-Out scheme:

$$J_{P,\{\beta_i\}} = \arg\min \left\{ \sum_{i=1}^{N} \left(\left\| Pz_i - PA_i\beta_i \right\|_F^2 + \lambda_1 \left\| \beta_i \right\|_1 \right) + \lambda_2 \left\| A - P^T PA \right\|_F^2 \right\}$$

$$\text{s.t. } PP^T = I$$

where β_i is the CR coefficient vector of z_i over A_i, λ_1 and λ_2 are scalar parameters.

The whole algorithm of the proposed fusing monogenic binary coding and collaborative representation is summarized as follows：

(1) When a face image comes, the MBC histogram is computed. Denote by X_i and y the dimensionality reduced features of MBC histogram of a training image and a testing image. The dictionary is denoted by X=[X_1, X_2,…,X_i].

(2) The projection matrix is computed as $P=(X^TX+\lambda I)^{-1}X^T$. y is coded over dictionary X by $\rho=Py$.

(3) The regularized residuals is computed as

$$r_i = \frac{\left\| y - X_i \hat{\rho}_i \right\|_2}{\left\| \hat{\rho}_i \right\|_2}$$

(4) The identify of y is output as Identity(y)=arg $\min_i\{r_i\}$.

5 Experimental Results and Analysis

In order to evaluate the effectiveness of the proposed method, extensive experiments were carried out on 3 standard databases: AR[17] PolyU-NIR[18] and CAS-PEAL[19]. All the experiments were carried out using Matlab version R2012a on a 3.30 GHz machine with 3.49 GB RAM.

In AR, as in [10], a subset (with only illumination and expression changes) that contains 50 male subjects and 50 female subjects was chosen from the AR dataset in our experiments. For each subject, the seven images from Session 1 were used for training, with the other seven images from Session 2 for testing. The images were cropped to 60×43. The comparison of competing methods is given in Table 1. In order to be consistent with other methods and provide a fair comparison, we used SDR reduce the dimension of each image to 300. In our test, 8 sampling points on a circle of radius of 8 are adopted, and a LBP or MBC image is divided into 5×5 sub-blocks. In the DR learning process, λ_1=0.03 and λ_2=1.5. LBP uses a nearest neighbor classifier and its similarity measure is based on the Chi-square distance. While cosine distance is used in MBC. The parameter λ in CRC is set as 0.01. Considering the accuracy and efficiency, we chose l1_ls[20] to solve the l1-regularized minimization in SRC.

Table 1. Recognition Rate and Speed on AR Database

Methods	Recognition rate /%	Time /s
NN	71.30	-
SVM	68.44	-
SRC	84.17	1.9633
CRC	84.69	0.0031
L-SRC [11]	93.29	2.0064
M-CRC	98.06	0.0579

In PolyU-NIR., The PolyU-NIR face database is a large scale near-infrared face database, consisting of 350 subjects, each subject providing about 100 samples. Various variations of face images, such as expression, pose, scale, focus, time, are involved in the capturing. In our experiments, a subset of 128 subjects (each subject providing about 50 samples) was chosen from the PolyU-NIR dataset. We randomly chose 5 samples per subject as training set, with the remaning as the testing set. The face images are normalized to 64×64 pixels. We used SDR reduce the dimension of each image to 300. In the DR learning process, λ_1=0.005 and λ_2=2.0. Other settings of the experiments are in accord with the testing on AR. The maximal recognition rate of each method is listed in Table 2.

Table 2. Recognition Rate and Speed on PolyU-NIR Database

Methods	Recognition rate /%	Time /s
NN	62.12	-
SVM	76.00	-
SRC	93.30	1.7878
CRC	93.70	0.0024
L-SRC [11]	94.75	1.8031
M-CRC	97.42	0.0401

In CAS-PEAL, the CAS-PEAL is a large-scale database of face photographs, which is constructed by the Joint R&D Laboratory for Advanced Computer and Communication Technologies (JDL) of Chinese Academy of Sciences (CAS). The CAS-PEAL face database contains 99,594 images of 1040 individuals (595 males and 445 females) with varying Pose, Expression, Accessory, and Lighting (PEAL). In this paper, we do three tests following the experiment setting in [21]. These three subsets are Expression (1200 training samples, 1040 gallery samples and 1884 testing samples), Accessory (1200 training samples, 1040 gallery samples and 2616 testing samples), and Distance (1200 training samples, 1040 gallery samples and 324 testing samples). The image is cropped to 100×100 by placing the two eyes at fixed locations. We used SDR reduce the dimension of each image to 300. In the DR learning process, λ_1=0.005 and λ_2=1.5. Other settings of the experiments are in accord with the testing on AR. The maximal recognition rate of each method is listed in Table 3.

Table 3. Recognition Rate and Speed on CAS-PEAL Database

Methods	Recognition rate /%			Time /s
	Expression	Accessory	Distance	
NN	53.70	37.10	74.20	-
SVM	57.32	38.98	69.83	-
SRC	65.66	55.73	93.06	2.3347
CRC	66.41	56.58	93.83	0.0051
L-SRC [11]	87.10	63.18	96.67	2.8648
M-CRC	91.51	68.90	98.46	0.5973

We can see that the proposed M-CRC outperforms other methods. However, it should be noticed that the computational burden of M-CRC is higher than the CRC. The performance of MBC is more competitive and even better than the state-of-art local feature such as Gabor, LBP, LGBP. The CRC can be seen a very powerful classifier. We can conclude that the proposed M-CRC could not only increase the discrimination of local features but also has powerful classification ability due to the use of CRC.

6 Conclusions

In this paper, we proposed a new approach based on Monogenic binary coding and collaborative representation based classification. When the MBC feature is extracted, the dimensionality problem is resolved by applying Sparse Dimensionality Reduction. And the CRC scheme is used to various pattern classification. The experimental results demonstrated that the proposed M-CRC is superior to other methods and has some potential to be applied in practical face recognition systems.

The future works include:

(1) Exploring other new and efficient local feature extraction scheme for the potential performance increase.

(2) Selecting more efficient scheme to reduce the feature dimension while enhancing its discrimination.

(3) Exploring more robust fusion scheme, which not only have significantly lower time and space complexity, but also have better recognition rates.

References

[1] Zhao W, Chellappa R, Phillips P J, et al. Face recognition: a literature survey [J]. *ACM Computing Survey*, 2003, 35(4): 399-458.

[2] Turk M, Pentland A. Eigenfaces for recognition [J]. *Cognitive Neuroscience*, 1991, 13(1):71-86.

[3] Belhumeur P N, Hespanha J P, Kriegman D J. Egienfaces vs. fisherfaces: recognition using class specific linear projection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(7): 711-720.

[4] Liu C, Wechsler H. Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition [J]. *IEEE TPAMI*, 2002, 11(4): 467-476.

[5] Ahonen T, Hadid A, Pietikainen M. Face recognition with local binary patterns [C]// *Proceedings of the 8th European Conference on Computer Vision, Prague*, 2004, LNCS3021: 469-481.

[6] Zhang W, Shan S, Gao W, et al. Local Gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition [C]// *Proceedings of IEEE International Conference on Computer Vision*, 2005, 1: 786-791.

[7] Yang M, Zhang L, Shiu S C K, et al. Monogenic binary coding: an efficient local feature extraction approach to face recognition [J]. *IEEE Trans. on Information Forensics and Security*, 2012, 7(6): 1738-1751.

[8] Heisele B, Ho P, Poggio T. Face recognition with support vector machine: global versus component-based approach [C]// *Proc. of IEEE Int'l Conf. Computer Vision*, 2001.

[9] Wright J, Yang A, Ganesh A, et al. Robust face recognition via sparse representation [J]. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2009, 31(2): 210-227.

[10] Zhang L, Yang M, Feng X. C. Sparse representation or collaborative representation: which helps face recognition? [C]// *In ICCV(2011)*, 471-478.

[11] Min R, Dugelay J. L. Improved combination of LBP and sparse representation based classification (SRC) for face recognition [C]// *Proceedings of IEEE Int'l Conf. Multimedia and Expo*, 2011.

[12] Kang C C, Liao S C, Xiang S.M, et al. Kernel sparse representation with local patterns for face recognition [C]// *Proceedings of IEEE Int'l Conf. Image Processing*, 2011.

[13] Zhang L, Yang M, Feng Z, et al. On the dimensionality reduction for sparse representation based face recognition [C]// *Proceeding of the ICPR*, 2010.

[14] Felsberg M, Sommer G. The monogenic signal [J]. *IEEE TSP*, 2001, 49(12): 3136-3144.

[15] Stein E M, Weiss G. *Introduction to Fourier Analysis on Euclidean Spaces. Priceton* [M]. NJ: Princeton Univ. Press, 1971.

[16] Zhang L, Yang M, Feng Z, et al. On the dimensionality reduction for sparse representation based face recognition [C]// *Proceeding of the ICPR*, 2010.

[17] Martinez A, Benavente R. *The AR Face Database* [R]. CVC Tech. Report No. 24, 1998.

[18] Zhang B C, Zhang L, Zhang D, et al. Directional binary code with application to PolyU near-infrared face database [J]. *Pattern Recognition Letters*, 2010, 31(14): 2337-2344.

[19] Gao W, Cao B, Shan S, et al. *The CAS-PEAL Larege-Scale Chinese Face Database and Evaluation Protocols* [R]. Technical Report JDL-TR-04-FR-001, Joint Research & Development Laboratory, CAS, 2004.

[20] Kim S J, Koh K, Lustig M, et al. A interior-point method for large-scale l1-regularized least squares [J]. *IEEE Journal on Selected Topics in Signal Processing*, 2007,1(4): 606-617.

[21] Shan S. G, Yang P, Chen X. L, et al. AdaBoost gabor fisher classifier for face recognition [C]// *Proceeding of IEEE International Workshop on Analysis and Modeling of Faces and Gestures (AMFG2005), Beijing, China*, LNCS 3723, 2005, 278-291.

FU Yu-xian is currently a M.S. degree candidate in College of Physics and Information Science from Hunan Normal University, Changsha, China. His research interests include image processing, pattern recognition, and especially focus on face recognition. He can be reached by 363533517@qq.com.

PENG Liang-yu is currently a Ph.D. and professor at College of Physics and Information Science from Hunan Normal University, Changsha, China. She received her B.S. degree at Xiangtan University, China, in 1986, and her M.S. and Ph.D. degree from Hunan University, in China, in 1995 and 2003 respectively. She was a postdoctor of Beijing University of Aeronautics and Astronautics. Her research interests include image processing and active filter design. She has authored over 100 papers in journals and conferences. She can be reached by liangyu_peng@163.com.