Adaptively Weighted Subpattern-Based Isometric Projection for Face Recognition

Lai Wei[1,⋆], Weiming Zeng[1], and Feife Xu[2]

[1] Department of Computer Science, Shanghai Maritime University, Haigang Avenue 1550, Shanghai, China
{weilai,wmzeng}@shmtu.edu.cn
[2] Department of Computer Science, Shanghai University of Electric Power, Pingliang Road 2103, Shanghai, China
xufeifei1983@hotmail.com

Abstract. In this paper, we propose an adaptively weighted subpattern-based isometric projection (Aw-spIsoP) algorithm for face recognition. Unlike IsoP (isometric projection) based on a whole image pattern, the proposed Aw-spIsoP method operates on sub-patterns partitioned from an original whole face image and separately extracts corresponding local sub-features from them. Moreover, the adjacency graph used in the algorithm is constructed based on path-based distance optimized neighborhoods of the sub-patterns and the contribution of each sub-pattern is adaptively computed in order to enhance the robustness to facial pose, expression and illumination variations. Experimental results on three bench mark face databases (ORL, YALE and PIE) show that Aw-spIsoP can overcome the shortcomings of the existed subpattern-based methods and achieve the promising recognition accuracy.

Keywords: face recognition, subpattern, path-based distance, isometric projection.

1 Introduction

Face recognition has been among the most active research topics in pattern recognition, computer vision and machine learning communities[1,2,3]. One of the most successful and well-studied techniques are the appearance-based methods[4,5,6]. Two of the most representative subspace techniques for face recognition are principal component analysis (PCA)[4] and fisher linear discriminant analysis (LDA)[5]. PCA is designed to reduce the dimension of the data by projecting the original data onto a linear subspace spanned by the leading eigenvectors of the data's covariance matrix. LDA searches for a low-dimensional subspace in which the data samples from the same class will assemble and the data samples with different class labels will lie apart.

Recently, a few new subspace techniques including locality preserving projection (LPP)[7], isometric projetion (IsoP)[8] and non-negative matrix factorization (NMF)[9], have attracted many researchers' attention. LPP is a manifold

⋆ Corresponding author.

H. Deng et al. (Eds.): AICI 2011, Part III, LNAI 7004, pp. 554–561, 2011.

learning based method which is effective in maintaining the locality of the face image data sets. Unlike LPP, IsoP aims to preserve the global manifold structure of the original data set. NMF is designed to capture part-based structures inherent in the face images space. The non-negative constraints of NMF do not allow negative elements either in the basis vectors or weighted vectors.

Subpattern-based algorithm[10,11,12] have also been proposed for face recognition. In the subpattern-based algorithms, face images are firstly divided into several small sub-images, and then the subpattern features are extracted by applying a given dimensionality reduction algorithm to all sub-image blocks. At last, faces are classified by comparing and combining the corresponding local features. A popular subpattern based technique is subpattern PCA (SpPCA)[13]. Besides PCA, some other techniques are also used for subpattern-based face recognition. Zhu proposed a subpattern non-negative matrix factorization (SpNMF)[14] algorithm. Wang et al. proposed an adaptively weighted subpattern-based LPP (Aw-spLPP) algorithm[11] and extended spLPP to a structure-preserved local matching approach (spLMA)[12].

In this paper, we propose an adaptively weighted subpattern-based isometric projection (Aw-spIsoP) algorithm for face recognition. In the proposed algorithm, the whole face images are also firstly partitioned into a set of equal-size sub-patterns in a non-overlapping way. Then IsoP algorithm is implemented on each of subpattern sets. Different from the traditional IsoP algorithm which uses Euclidean distance to find the k nearest neighbors for constructing the adjacent graph, we use path-based distance to determine the neighbors for each sample. In succession, the contribution of each sub-pattern to recognition is adaptively computed using path-based distance and class information. Finally, subpattern-based features are compared and combined for face recognition. Experimental results on the bench mark face databases show that the Aw-spIsoP algorithm achieves the promising recognition accuracy.

2 Review of Isometric Projection

Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n] \in \mathbb{R}^{D\times n}$ be the data matrix, $d_{\mathcal{M}}$ be the geodesic distance measure on the data manifold $\mathcal{M}$ and d_e the standard Euclidean distance measure in $\mathbb{R}^d$. Isometric Projection aims to find a embedding function f such that Euclidean distances in $\mathbb{R}^d$ can provide a good approximation to the geodesic distances on $\mathcal{M}$. That is,

$$f = \arg\min_f \sum_{i,j} \Big(d_{\mathcal{M}}\big(\mathbf{x}_i - \mathbf{x}_j\big) - d_e\big(f(\mathbf{x}_i) - f(\mathbf{x}_j)\big)\Big)^2 \tag{1}$$

The geodesic distances $d_{\mathcal{M}}(i,j)$ between all pairs of points on the manifold $\mathcal{M}$ can be estimated by computing their shortest path distances $d_{\mathcal{G}}(i,j)$ on the adjacency graph G modeling the local geometry of the data set. The graph usually is constructed by KNN (k-nearest-neighbors) or $\epsilon - ball$ method.

Let D be the distance matrix such that D_{ij} is the distance between $\mathbf{x}_i$ and $\mathbf{x}_j$. Define matrix $S_{ij} = D_{ij}^2$ and $H = I - \frac{1}{m}\mathbf{e}\mathbf{e}^T$, where I is the identity matrix and

$\mathbf{e}$ is the vector of all ones. It can be shown that $\tau(D) = -HSH/2$ is the inner product matrix. Let D_Y denote the Euclidean distance matrix in the reduced subspace, and $\tau(D_Y)$ be the corresponding inner product matrix. Thus, the objective function (1) becomes minimizing the following:

$$\|\tau(D_G) - \tau(D_Y)\|_{L^2} \tag{2}$$

where $\|\cdot\|$ is the L^2 matrix norm.

Consider a linear function $Y = (\mathbf{y}_1, ..., \mathbf{y}_n) \in \mathbb{R}^{d\times n} = \mathbf{w}^T X$, d is the dimensions of embedding subspace. Then $\tau(D_Y) = Y^T Y = X^T \mathbf{w}\mathbf{w}^T X$. According to function (2), the optimal projection is given by solving the following minimization problem:

$$\mathbf{w}^* = \min_{\mathbf{w}} \|\tau(D_G) - X^T \mathbf{w}\mathbf{w}^T X\| \tag{3}$$

To avoid degenerate solutions, an additional constrain $\mathbf{w}^T X X^T \mathbf{w} = 1$ should be imposed on the above problem. With simple algebric formulation, the optimization problem can be expressed as follows:

$$\arg \max_{\mathbf{w}^T X X^T \mathbf{w}=1} \mathbf{w}^T X \tau(D_G) X^T \mathbf{w} \tag{4}$$

3 Adaptively Weighted Isometric Projection

The proposed algorithm consists of three main steps: (1) partition face images into sub-patterns, (2) apply IsoP to local feature extraction and compute the contribution of each subpattern (3) classify an unknown face image.

3.1 Face Image Partition

In subpattern-based face recognition methods, a face image can be partitioned into a set of equally or unequally sized sub-images. Without loss of generality, equally sized partition is adopted in our approach as in many other approaches(such as Aw-spPCA, Aw-spLPP, spNMF and spLMA)[11,12,13,14].

Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n]$ denote n face images belonging to n_p persons, and the size of each image is $H_1 \times H_2$. We first partition each face image into C equally sized sub-images in a non-overlapping way, and further concatenate them into corresponding column vectors with dimensionality of $H_1 \times H_2/C$. Afterwards, the sub-pattern vectors at the same position of all face images are collected to form a specific subpattern set. Therefore, we can get C separate subpattern sets totally. The procedure of image partition is illustrated in Figure 1.

3.2 Subpattern-Based IsoP and Contributions Computation

A.Subpattern-based IsoP. In the traditional IsoP algorithm, Euclidean distance is used to find the k nearest neighbors of each sample. In this paper, we will

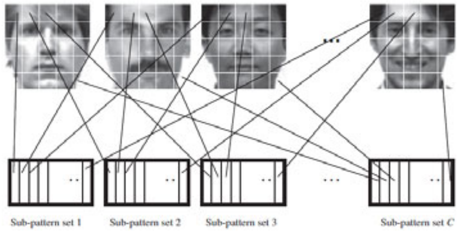

Fig. 1. The construction of sub-image sets[13]

use path-based distance[15,16] to build the neighborhood graph for Aw-spIsoP. We briefly introduce the definition of path-based distance as follows.

Let $G = (V, V \times V)$ be the complete Euclidean graph of all data points. And P_{ij} denote the set of all paths from vertex i to vertex j through the graph. The path-based distance $d_p(v_i, v_j)$ can be defined as follows:

$$d_p(v_i, v_j) = \min_{\mathfrak{p} \in P_{ij}} \{ \max_{1 \leq h \leq |\mathfrak{p}|} d_e(v_h, v_{h+1}) \} \tag{5}$$

Then for each subpattern set $S_m (m = 1, 2, \cdots, C)$, its low-dimensional features can be extracted by IsoP with path-based distance. Let $Z_m = [\mathbf{z}_{m1}, \mathbf{z}_{m2}, \cdots, \mathbf{z}_{mn}]$ denote n column vectors in S_m. According to the description in Section 2, the projection matrix $\mathbf{W}_m$ of S_m can be composed of the eigenvectors corresponding to its largest d eigenvalues of the following generalized eigenvalue problems:

$$Z_m \tau(D_{mG}) Z_m^T \mathbf{w}_{mj} = \lambda Z_m Z_m^T \mathbf{w}_{mj} \tag{6}$$

where $j = 1, 2, \cdots, d$. $\mathbf{W}_m = [\mathbf{w}_{m1}, \mathbf{w}_{m2}, \cdots, \mathbf{w}_{md}]^T$. Then the low-dimensional embedding of $\mathbf{z}_{m1}$ can be achieved as $\mathbf{y}_{m1} = \mathbf{W}_m^T \mathbf{z}_{m1}$.

Contributions computation. In our method, the contribution of each subpattern to recognition is computed through the label information of one subpattern and its k nearest neighbors chosen by path-based distance.

Suppose $\mathbf{z}_{mi}$ is the mth subpattern of the ith image $\mathbf{x}_i$. Then the weight of the mth sub-pattern can be obtained as

$$E_m = \frac{1}{n} \sum_{i=1}^{n} \frac{\sum_j exp(d_p(\mathbf{z}_{mi}, \mathbf{z}_{mj})/t)}{\sum_l exp(d_p(\mathbf{z}_{mi}, \mathbf{z}_{ml})/t)} \tag{7}$$

where $j, l \in \mathcal{N}(\mathbf{z}_{mi})$, $\mathcal{N}(\mathbf{z}_{mi})$ is the neighborhood of $\mathbf{z}_{mi}$. t is a positive parameter(in the following experiments, t is empirically set as the mean of the all pairwise distance of the data points). $\mathbf{x}_j$, $\mathbf{x}_i$ should be the face images of same person. Obviously, we have $0 \leq E_m \leq 1$.

3.3 Face Recognition

For the sake of classifying an unknown face image, the unknown face image x^* should be firstly divided into C sub-images in the same way previously applied to the training images. Then, each unknown subpattern's features are extracted by using the corresponding projection matrix $\mathbf{W}_i(i = 1, 2, \cdots, C)$. The class label of each sub-pattern is determined by a nearest neighbor classifier using Euclidean distance. Since one classification result for the unknown sample is generated independently in each subpattern, there will be total C results from C sub-patterns. To combine C classification results from all sub-patterns of this face image, a weighted voting method is used. Let the probability of the unknown image x^* belonging to the cth class be

$$p_c = \sum_{i=1}^{C} E_i q_i^c \tag{8}$$

where

$$q_i^c = \begin{cases} 1, \text{if the } i\text{th sub-pattern belongs to } c\text{thclass;} \\ 0, \text{otherwise.} \end{cases} \tag{9}$$

The final classification result is

$$L(\mathbf{x}^*) = \arg\max_c(p_c) \tag{10}$$

4 Experiments

In this section, the performance of the proposed algorithm (Aw-spIsoP) will be explored systematically on three standard face databases(ORL, YALE and PIE). The information of each data set is briefly introduced as follows.

ORL[1] database contains 10 different images of each of 40 distinct subjects. For some subjects, the images were taken at diffrent times, varying the lighting, facial expressions. YALE[2] database contains 165 grayscale images of 15 individuals. There are 11 images per subject, one per different facial expression or configuration. The CMU PIE[3] face database contains 68 subjects with 41,368 face images as a whole. The face images were captured by 13 synchronized cameras and 21 flashes, under varying pose, illumination, and expression. We only use a subset containing 5 near frontal poses (C05, C07, C09, C27, C29) and all the images under different illuminations and expressions. In our experiments, all the face images are resized to 64×64. The sample images of these three data sets are shown in Figure 3.

Recognition accuracy is used to evaluate the algorithm's performance. For comparisons, the seven different algorithms that we evaluated are Aw-SpPCA,

[1] http://www.uk.research.att.com/facedatabase.html
[2] http://cvc.yale.edu/projects/yalefaces/yalefaces.html
[3] http://www.ri.cmu.edu/projects/project_418.html

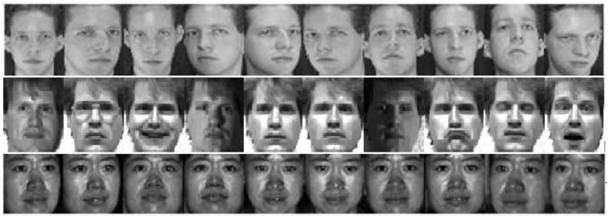

Fig. 2. The sample images of different face datasets. The images in the first, second and third rows are sample images of ORL database, YALE database and CMU PIE database respectively.

Aw-spLPP(the kernel width is empirically set as the mean of all pairwise distance of the data points), spNMF, IsoP, PCA, supervised LPP (SLPP, namely the affinity graph is constructed using label information of the training samples) and NMF. We first randomly select a certain number of face images from each person for each data set. Those selected images will be used as training samples and the remaining images will be used for testing. To control the balance of computation cost and recognition accuracy, the size of sub-image is set to 16×8 for subpattern-based algorithms. The experiments are conducted over 10 times independent runs.

A. Experimental results using ORL dataset. The recognition results on ORL dataset versus subspace dimensions are shown in Figure 4. Figure 4(a) and 4(b) show the performance of different algorithms with 4 and 5 training samples respectively. In these experiments, the neighbor parameters in Aw-spIsop, IsoP and Aw-spLPP are set as 4.

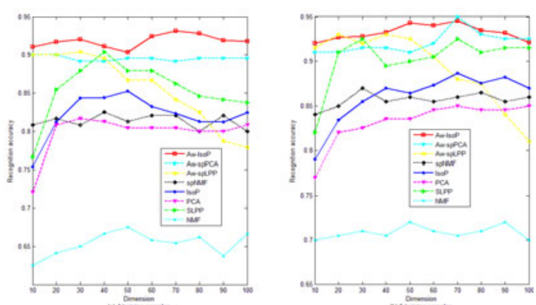

Fig. 3. Performance comparisons of different algorithms on ORL database

B. Experimental results using YALE dataset. The recognition results on YALE dataset versus subspace dimensions are shown in Figure 5. Figure 5(a) and 5(b) show the performance of different algorithms with 5 and 6 training samples

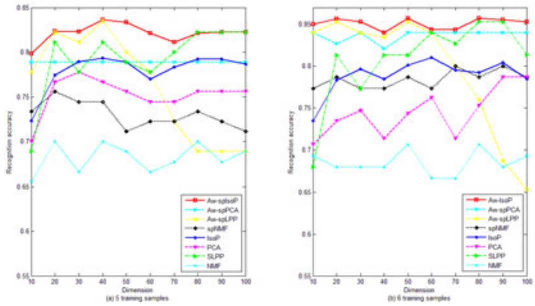

Fig. 4. Performance comparisons of different algorithms on YALE database

respectively. The neighbor parameters in Aw-IsoP, Isop and Aw-spLPP are set as 5.

C. Experimental results using PIE dataset. We randomly select 6 images for each individual on each pose (C05, C07, C09, C27, C29) to build up a new data set (denotes as S-PIE) used in the following experiments. Hence, there are 30 images for each individual in S-PIE. The performance of different algorithms with 10 and 15 training samples are evaluated. The experimental results are shown in Figure 6. The neighbor parameters in Aw-IsoP, Isop and Aw-spLPP are set as 8.

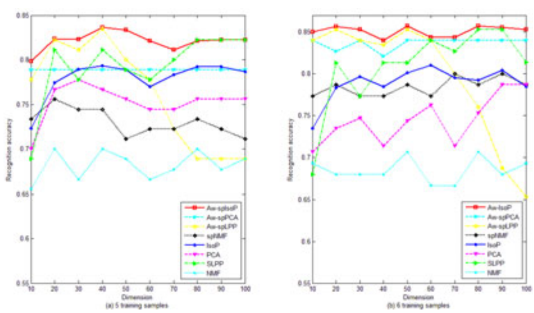

Fig. 5. Performance comparisons of different algorithms on S-PIE database

5 Conclusion

In this paper, we propose a new subpattern-based face recognition algorithm, called adaptively weighted subpattern-based isometric projection (Aw-spIsop). Aw-spIsoP uses path-based distance to construct the adjacency graph and computes the contribution of each subpattern adaptively with class information and path-based distance. We perform experiments on three face image databases

(ORL ,YALE and PIE). The experimental results indicate that our proposed approach is effective.

References

1. Pantic, M., Rothkrantz, L.: Automatic analysis of facial expressions: the state of art. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(12), 1424–1445 (2000)
2. Fasel, B., Luettin, J.: Automatic facial expression analysis: a survey. Pattern Recognition 36, 259–275 (2003)
3. Tan, X., Chen, S., Zhou, Z., Zhang, F.: Face recognition from a single image per person: A survey. Pattern Recognition 39, 1725–1745 (2006)
4. Turk, M., Pentland, A.: Face recognition using eigenfaces. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 586–591 (1991)
5. Belhumeur, P., Hespanha, J., Kriegman, D.: Eigenfaces vs. fisherfaces: recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence 19(7), 711–720 (1997)
6. Levin, A., Shashua, A.: Principal Component Analysis over Continuous Subspaces and Intersection of Half-Spaces. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002. LNCS, vol. 2352, pp. 635–650. Springer, Heidelberg (2002)
7. He, X., Yan, S., Hu, T., Niyogi, P., Zhang, H.: Face recognition using Laplacianfaces. IEEE Transactions on Pattern Analysis and Machine Intelligence 27(3), 328–340 (2005)
8. Cai, D., He, X., Han, J.: Isometric Projection. In: Proc. 22nd Conference on Artifical Intelligence (AAAI), Vancouver, Canada (July 2007)
9. Lee, D., Seung, H.: Algorithms for non-negative matrix factorization. In: Advances in Neural Information Process, pp. 556–562 (2000)
10. Zou, J., Ji, Q., Nagy, G.: A comparative study of local matching approach for face recognition. IEEE Trans. Image Process 16(10), 2617–2628 (2007)
11. Wang, J., Zhang, B., Wang, S., Qi, M., Kong, J.: An adaptively weighted subpattern locality preserving projection for face recognition. Journal of Network and Computer Applications 33, 323–332 (2010)
12. Wang, J., Zhang, B., Wang, S., Qi, M., Kong, J.: A structure-preserved local matching approach for face recognition. Pattern Recognition Letters 32, 494–504 (2011)
13. Chen, S., Zhu, Y.: Subpattern-based principle component analysis. Pattern Recognition 37, 1081–1083 (2005)
14. Zhu, Y.: Sub-pattern non-negative matrix factorization based on random subspace for face recognition. In: International Conference on Wavelet Analysis and Pattern Recognition, pp. 1356–1360 (2007)
15. Fischer, B., Buhmann, J.: Path-based clustering for grouping of smooth curves and texture segmentation. IEEE Trans. Pattern Analysis and Machine Intelligence 25(4), 513–518 (2003)
16. Wen, G., Jiang, L., Shadbolt, N.: Using Graph Algebra to Optimize Neighborhood for Isometric Mapping, In:Veloso MM,ed. In: Veloso, M.M. (ed.) Proc.of the 20th Int'l Joint Conf. on Artificial Intelligence (IJCAI 2007), pp. 2398–2403 (2007)