A Pose Robust Face Recognition Approach by Combining PCA-ASIFT and SSIM

Wei Qi, Yaxi Hou, Lifang Wu*, and Xiao Xu

School of Electronic Information and Control Engineering,
Beijing University of Technology, Beijing, 100124, China
lfwu@bjut.edu.cn,qiwei7@sina.com,
{xuxiao2013,houyaxi}@emails.bjut.edu.cn

Abstract. Affine Scale Invariant Feature Transform (ASIFT) is robust to scales, rotation, scaling and affine transformation. It could be used for face recognition with pose variation. However, ASIFT requires large data. Could we reduce the data of ASIFT and preserve the face recognition performance? In this paper, we propose an effective face recognition algorithm to combining the structural similarity (SSIM) and PCA-ASIFT (PCA-ASIFT&SSIM).First, we reduce ASIFT dimension using principal component analysis and get PCA-ASIFT. The PCA-ASIFT's discriminative capability drops because of the dimension reduction. It brings about more false SIFT matching. We further introduce the SSIM to reduce the false matching. The experimental results show the efficiency of the proposed approach.

Keywords: Face recognition, dimension reduction, PCA-ASIFT, SSIM.

1 Introduction

Face recognition is one of the active research topics in computer vision and pattern recognition. Recently, face recognition technology has reached a good performance under a controlled imaging condition. But the performance of face recognition with pose variation drops significantly. Scale invariant feature (SIFT) [2] is the local feature with rotating invariance and good robustness. Affine-SIFT (ASIFT) [3] is further proposed to improve the robustness to affine transformation. ASIFT can solve the problem of face recognition under different poses much better.

WU et. Al. [4] proposed a face recognition approach to combining ASIFT and SSIM[4], but they endured heavy storage cost. They obtain roughly 3000 ASIFT points in a face image of 80×80. Each ASIFT point has 128dimensiondescriptors. If each dimension is represented by one byte, the data for all ASIFT features from one face image is about $3000 \times 128 = 384000$bytes. It is much larger than an 80×80 face's image data size: $80 \times 80 \times 3 = 19200$ bytes. Therefore, it is necessary to reduce the data. In this paper, dimension reduction is studied.

High-dimensional data is multivariate data obtained by observing. It describes from different angles or method of the same object. Obviously, with the increasing of

* Corresponding author.

Z. Sun et al. (Eds.): CCBR 2014, LNCS 8833, pp. 163–172, 2014.

data dimension, it provides richer and more detailed information with the emergence of the "curse of dimension" [5].The concept proposed by the Bellman, for the known number of samples, there is a maximum value of the number of feature, when the actual number exceeds the maximum; the performance of the classifier is not improved but degenerated. In order to solve the dimension disaster, dimension reduction is proposed. The basic principle is that the sample points are mapped from the input space to a lower dimensional spacesoastoobtain a compact, low-dimensional representative.

Depending on the mapping method, dimension reduction methods are divided into being linear and being nonlinear. Linear dimension reduction methods include principal component analysis (PCA), projection pursuit (PP), linearly discriminant analysis (LDA), locality preserving projections (LPP),sparsity preservingprojections (SPP) and so on. The principle of linear dimension reduction is to find a linear projection model in high-dimensional dataspace. It has poor effect on nonlinear high-dimensional data structure. Thus, methods of nonlinear dimension reduction are proposed, like multidimensional scaling (MDS), ISOMAP, KernelBasedPrincipal Component Analysis (KPCA), locally linear embedding (LLE) and so on.

Principal Component analysis (PCA) is to project the high-dimensional data into a low dimension while keeping its spatial features as much as possible. It is a popular dimension reduction method and is usually used in image processing. Therefore, PCA is used in this paper. In our experiments, AISFT feature dimension could be reduced using the PCA, however, it also brings the loss of information and the performance of face recognition with the point that pose variation can't reach the performance of the ASIFT method. Structural similarity (SSIM) is a method used to evaluate the image quality. It is a structural measure based on the similarity between pixels. It can be applied to face recognition and to analyze the similarity of the two facial images. We further introduce SSIM to improve the face recognition performance.

We propose a face recognition approach to combining PCA-ASIFT and SSIM. Firstly, PCA is utilized to reduce the dimension of ASIFT descriptors, and we get PCA-ASIFT. At face recognition stage, the PCA-ASIFT is matched firstly, then for each PCA-ASIFT matching points, its SSIM is used to further filter out the mismatched points, and the face authentication is finished by the average SSIM.

2 The Proposed Approach

We propose a framework of face recognition combining SSIM and PCA-ASIFT. As shown in Fig. 1, we first extract AISFT descriptors from face image, then we utilize principal component analysis (PCA) to reduce AISFT descriptors dimension and obtain a new kind of descriptor with low dimension. We define it PCA-ASIFT descriptors. The method to obtain PCA-ASIFT descriptors is defined as PCA-ASIFT algorithm. For a probe face image, the similarity between it and the PCA-ASIFT of each subject are measured. We do preliminary face features' matching and get pairs of matching points. Then, we utilize structural similarity algorithm to compute the similarity of pairs of local images which is also called face patches. The face patch is window pixels which matching point as the center. The size of window is set to 5×5. After computing the SSIM index, we retain pairs of matching points which are above the threshold. The threshold is set by analyzing the distribution of the SSIM index.

Then we obtain the real matching points. Face recognition is implemented by judge whether the amount of the matching points of the subject is the largest one.

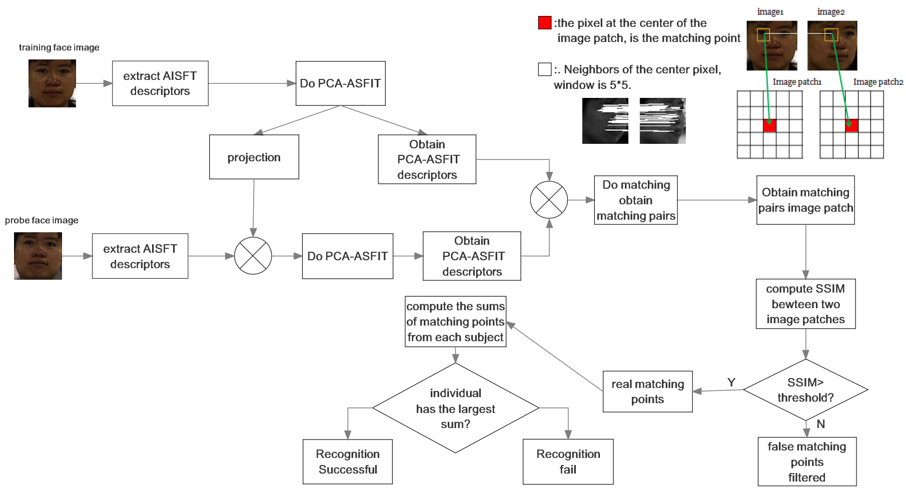


Fig. 1. The flowchart of proposed framework

3 PCA-ASIFT

3.1 Extraction of PCA-ASIFT

PCA-ASIFT uses principal component analysis (PCA) to reduce 128 dimension of ASIFT descriptor. We define an original ASIFT matrix which is composed of all the ASIFT vectors extracted from a face image. We get n ASIFT points from one face image, in which i^{th} ASIFT point is denoted as $\mathrm{D_i}$. Each ASIFT point is 128 dimensional descriptor, so $D_i = \{f_{i1}, f_{i2}, \ldots, f_{ij}, \ldots, f_{i128}\}$.And the original matrix can be denoted as $\{D_1, D_2, \ldots, D_n\}$T ,the size of the matrix is n × 128.

Then we do PCA on the original ASIFT matrix. The details are as follows:

1. Normalize the matrix, each column subtracts corresponding average and divide by variance.
2. Get the covariance matrix R.
3. UtilizeJacobialgorithm [8] to calculate the eigenvalues and eigenvectors of matrix R.
4. Put eigenvalues into descending order and get new order of the eigenvalues. Then we use K eigenvalues which is on the order as principal component to form a new matrix. The matrix is defined as projection matrix P.
5. Get the PCA-ASIFT matrix Y, Y is denoted as Y=X×P, the size of the matrix is n×k, k is the dimension of PCA-ASIFT vector after reducing by PCA.

3.2 PCA-ASIFT Matching Based on Euclidean Distance

We measure the distinctiveness of features between training face image and a probe face image. First, we extract ASIFT points from a probe face image. Then, these AISFT points compose a matrix. The following is to normalize the matrix and project the matrix to the same space which PCA-ASIFT points of the training face image belong to. After we get the PCA-ASIFT points from the probe face image, we do PCA-ASIFT matching. The PCA-ASIFT matching still uses ASIFT matching method[4]. Each PCA-ASIFT points from probe face image should calculate Euclidean distance with all the PCA-ASIFT points one by one. Then we find the minimum Euclidean distance of the two feature points and the second minimum to calculate the relative ratio. If the ratio is less than the threshold ratio, we consider that the two features from probe and training face image are matching.

4 Structural Similarity (SSIM)

4.1 Definition of SSIM

Structural similarity (SSIM) is a full-reference image quality evaluation method proposed by Zhou Wang et al [6]. Structural similarity theory suggests that there is a strong correlation between the pixels and the highly structured image. Considering the generation of image, structural information reflects the structure of objects in the scene. It should be independent of the brightness and the contrast. Structural similarity constituted by three factors: brightness, contrast, and structure similarity. Combining the similarity of these three, constitute the reference image x and distorted image y similarity operator, the formula is as follows:

$$SSIM(x,y) = [l(x,y)]^{\alpha} \times [c(x,y)]^{\beta} \times [s(x,y)]^{\gamma} \tag{1}$$

Where, *l(x, y), c(x, y) and s(x, y)* are the brightness contrast function, the contrastcontrast function and the structural contrast function respectively.

$$\begin{cases} l(x,y) = \dfrac{2\mu_x\mu_y + C_1}{{\mu_x}^2 + {\mu_y}^2 + C_1} \\ c(x,y) = \dfrac{2\sigma_x\sigma_y + C_2}{{\sigma_x}^2 + {\sigma_y}^2 + C_2} \\ s(x,y) = \dfrac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \end{cases} \tag{2}$$

Where μ is the average of image x, y, reflects the brightness information. σ is the variance of image x, y, reflect the contrast information. σ_{xy} is the covariance of image x and y, reflect the structure information.

SSIM index has the following character:

(1) symmetry : *SSIM* (*x, y*) = *SSIM* (*y, x*)

(2) $SSIM \leq 1$

(3) If and only if *x*=*y*, *SSIM*=1.

4.2 Analysis of SSIM at PCA-ASIFT Points

We utilize structural similarity (SSIM) to filtrate some matching points that are lowly similar after dimension reduction. And we propose SSIM to improve the correct rate for face recognition. Ideally, the value of SSIM between correct matching point and false matching point are obviously different. The ideal state is the distribution of SSIM value is distinct from 0 to 1. In this paper, we set the size of window as 5×5. Fig.2 shows the distribution of SSIM value with the window size of 5×5. And it can distinguish within class pairs from between class pairs.

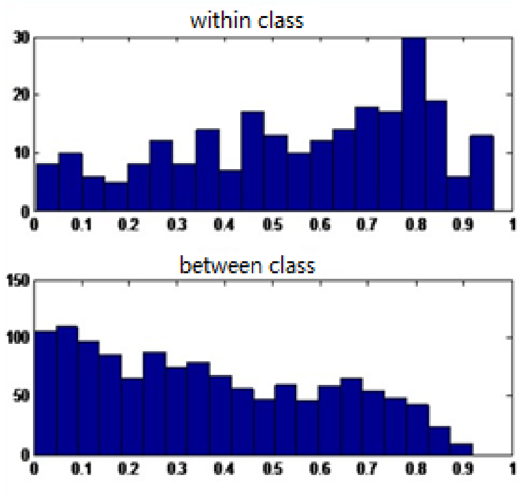


Fig. 2. The distribution of SSIM value with the window size is 5×5 from within and between class pairs

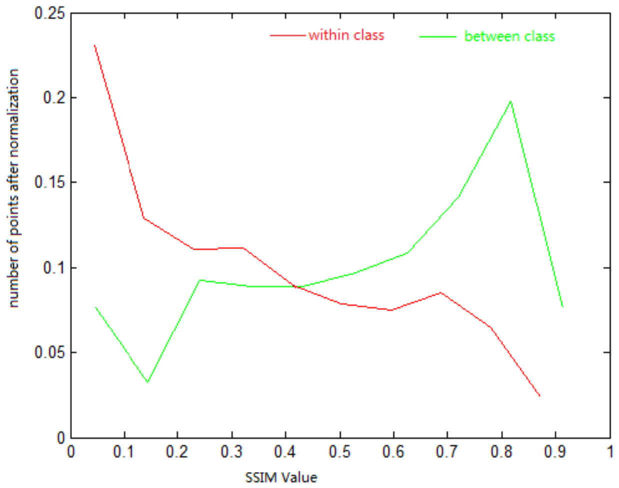


Fig. 3. The distribution of SSIM value within and between class pairs

Why can we use SSIM to improve the performance of face recognition? Fig.3 shows the distributionof SSIM value between within and between class pairs after normalization. We assume the matching pairs from the same subject as within class pairs, other pairs as between class. And we normalize the distribution by the total number so that we could distinguish the matching pairs from the same subject easily. It is obviously observed that the SSIMs of most within class pairs are bigger than 0.7, while those of between class pairs are smaller than 0.2. It means that they could be discriminated. Therefore, SSIM could be used for further improvement work.

4.3 SSIM Computation for Face Recognition

We extract ASIFT points from the face image and use PCA to do dimension reduction. Then we get PCA-ASIFT points and do match to get match points. SSIM's combination with PCA-ASIFT is for further analysis of the similarity between the matching points pairs. In essence, it is for seeking local image structural similarity of which matching points is the center and we set an appropriate threshold to throw some less-similar pairs of matching points.

SSIM is computed as follows:

1. Matching still use 3.2 PCA-ASIFT matching method. We save the location of the PCA-ASIFT.
2. Select the appropriate size of the window, holding match point as the center and looking for neighborhood pixels. If you choose the pixel neighborhood over the boundaries of the image, edge processing should be done. Definite beyond the boundary points with adjacent positions boundary pixels instead.
3. Because the input is bmp format images, we do RGB components for SSIM respectively, and we sum three SSIM values, and then divide the sum by 3. The result is regarded as the final structural similarity value. The first step to do SSIM is to calculate the average, variance, covariance of the pixels within the window.
4. Compare the final SSIM values with thresholds. If the final SSIM value is greater than the threshold, which we think the matching point is correct one, if it is not, the matching points will be filtered.

We compute SSIM for each pair of the matched PCA-ASIFT points. If SSIM is greater than the threshold, this pair is preserved. Otherwise, it should be discarded. Face recognition is implemented by the number of matched pairs.

5 Experiments

We test our approach on the CMU-PIE [1] face database sand the Extended Yale Face Database B(Extended YaleB) [10]. They are commonly used databases for face recognition across pose variation.

The CMU-PIE face database includes 68 individuals. We choose face images with 5 pose for each individual. They are face images of front view (0° (C27)), up/down rotation 30° (C09/C07) and left/right rotation 25° (C05/C29), as shown in Fig.4.

Fig. 4. Example face images from CMU-PIE

The Extended Yale Face Database B (Extended YaleB) includes 28 individuals. The data format of this database is the same as the Yale Face Database B [9].Fig.5 shows the face images of the same individual. Face images of front view (0° (P0)) is selected as training face image. The face images of up/down rotation 12°(P1/P5) ,both left and up/down rotation 12°(P2/P4)and left rotation 12°(P3) are selected as probe ones.

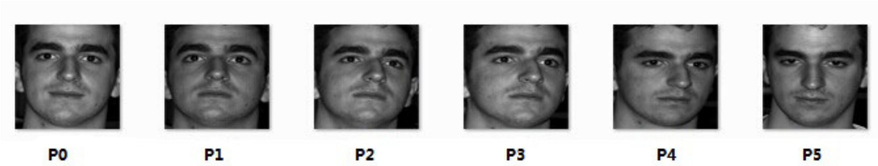


Fig. 5. Example face images from Extended YaleB

5.1 Comparison of Different Dimensions

The dimension of ASIFT is reduced to 96, 64and 32 respectively. When PCA algorithm extract main components to generate a new projection space, rateofcontribution S is often considered as standard. Formula 3 shows how to calculate the rateofcontribution S.

$$S = \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{m} \lambda_i} \tag{3}$$

Where, $_i$ is the eigenvalue. When the rate of contribution reaches the standard, we retain features which arecurrently component to compose the low-dimensional projection matrix. The PCA-ASIFT algorithms calculate the rate of contribution for different dimensions is 96 dimensions is roughly 85% -90%, 64dimensions is roughly 65% -70% and 32 dimensions is roughly 40% -45%.

Table 1 shows that we got greater performance after combining with SSIM on both CMU-PIE and extendingYaleB. After we do PCA-ASIFT, we can find that the lower dimension is, the recognition rate gets lower. However, the recognition rate is higher after SSIM combines with PCA-ASIFT and the effect of dimension is not obvious on the two databases. We suggest using PCA-ASIFT which can reduce the dimension of ASIFT descriptor into 64.

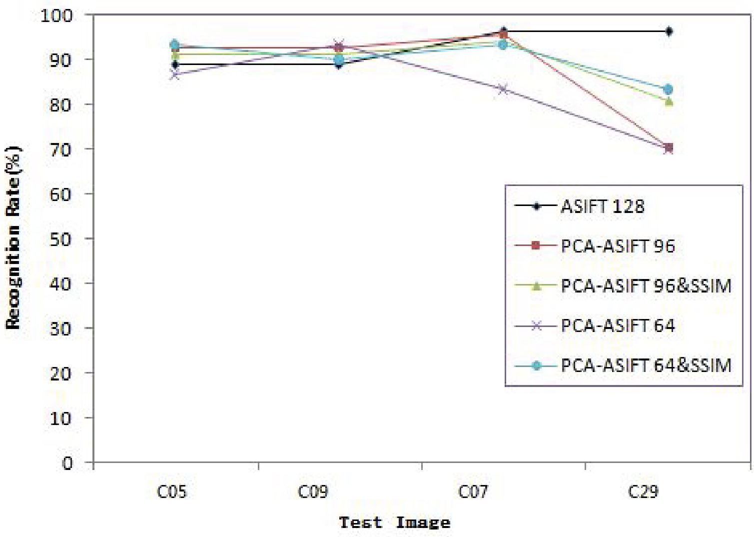
Table 1. The Recognition Rates with Different Dimension Before and After do SSIM

Method	PCA-ASIFT Dimension	Average On CMU-PIE database	Average On extended YaleB database
PCA-ASIFT	96	87.9%	92.1%
PCA-ASIFT	64	83.3%	89.3%
PCA-ASIFT	32	70.0%	78.6%
PCA-ASIFT&SSIM	96	89.3%	96.4%
PCA-ASIFT&SSIM	64	90%	97.1%
PCA-ASIFT&SSIM	32	81.3%	92.9%

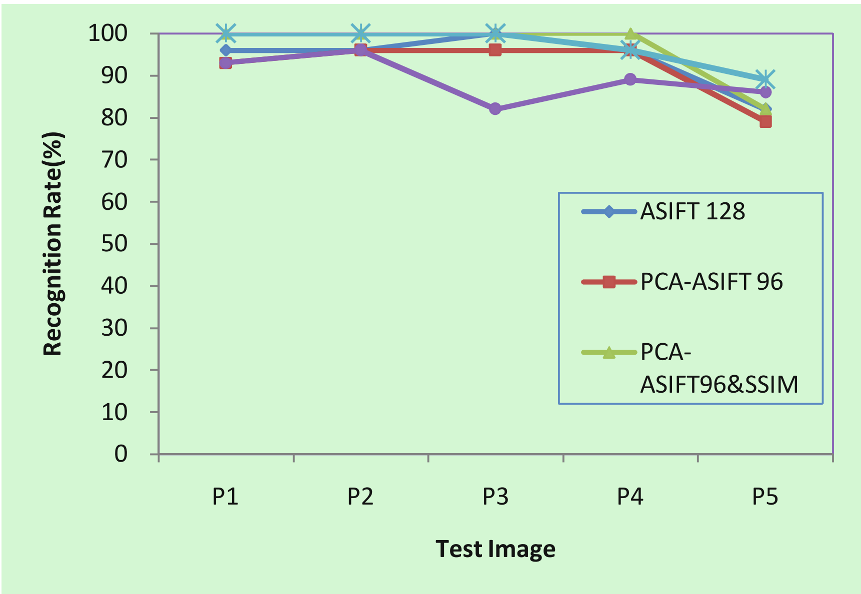
5.2 Comparison of Different Algorithms

We further compare our algorithms of PCA-ASIFT with dimension of 96(PCA-ASIFT 96), dimension of 64(PCA-ASIFT 64), PCA-ASIFT 96 & SSIM and PCA-ASIFT 64 &SSIM with the original ASIFT on two databases. The compared results are shown in Fig.6 (a). Fig.6 (a) demonstrates that the performance of PCA-ASIFT 96's drops with C29, 64's drops with C29 and C07compare with original ASIFT. However, we can observe that PCA-ASIFT 64& SSIM improves the performance a lot from Fig.6(a). And the original ASFIT gets the best performance. For the other two test face images, the three algorithms get good performance. Fig.6 (b) shows that great performance of PCA-ASIFT&SSIM, both PCA-ASIFT96 &SSIM and PCA-ASIFT64 &SSIMachieve 100% recognition rates with some poses on extended YaleB database. Comparing with the performance of original ASIFT, PCA-ASIFT 96's and 64's drops little with last three poses (P3, P4and P5). After combines with SSIM, we can obviously see that both PCA-ASIFT96 &SSIM and PCA-ASIFT64&SSIMperformance are improved and are better than original ASIFT algorithm.

Our approach PCA-ASIFT & SSIMachieve great performance and is better than the original ASIFT algorithm and the algorithm of only using PCA-ASIFT to reduce the dimension. We suggest using PCA-ASIFT to reduce the dimension into 64 and then combine with SSIM.



(a)CMU-PIE database



(b)Extended YaleB database

Fig. 6. The performance of different method under controlled imaging condition on two database

6 Conclusion

In this paper, we focus on the facial features ASIFT dimension reduction. We propose PCA-ASIFT & SSIM for dimension reduction with the preserved performance. The experimental results show that our PCA-ASIFT method is robust to controlled imaging condition of within off-plane rotation of 25 degree. In the future, we will extend this work to face images of other poses, so that our approach is robust to pose variation.

References

1. Sim, T., Baker, S.: The CMU Pose, Illumination Expression Database. IEEE PAMI 25(12), 1615–1618 (2003)
2. Lowe, D.G.: Distinctive image features from scale-invariant key points. IJCV 60(2), 91–110 (2004)
3. Jean-Michel, M., Yu, G.: ASIFT: A New Framework for Fully Affine Invariant Image Comparison. SIAM Journal on Imaging Sciences 2(2), 438–468 (2009)
4. Wu, L., Zhou, P., Liu, S., Zhang, X., Trucco, E.: A Face Authentication Scheme Based on Affine-SIFT (ASIFT) and Structural Similarity (SSIM). In: Zheng, W.-S., Sun, Z., Wang, Y., Chen, X., Yuen, P.C., Lai, J. (eds.) CCBR 2012. LNCS, vol. 7701, pp. 25–32. Springer, Heidelberg (2012)
5. Yu, G., Morel, J.-M.: A fully affine invariant image comparison method. In: ICASSP, pp. 1597–1600 (2009)
6. Bellman, R., Kalaba, R.: Adaptive Control Processes. IRE Transactions on Automatic Control 4(2), 1–9 (1959)
7. Wang, Z., Bovik, A.C., Sheikh, H.R., et al.: Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing 13(4), 600–612 (2004)
8. Golub, G.H., van der Vorst, H.A.: Eigenvalue computation in the 20th century. Journal of Computational and Applied Mathematics 123(1-2), 35–65
9. Georghiades, A., Belhumeur, P., Kriegman's, D.: From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose. PAMI (2001)
10. Lee, K.-C., Ho, J., Kriegman, D.: Acquiring Linear Subspaces for Face Recognition under Variable Lighting. PAMI (May 2005)