JOURNAL OF INFORMATION PROCESSING SYSTEMS JIPS

Viewpoint Unconstrained Face Recognition Based on Affine Local Descriptors and Probabilistic Similarity

Yongbin Gao* and Hyo Jong Lee*,**

Abstract

Face recognition under controlled settings, such as limited viewpoint and illumination change, can achieve good performance nowadays. However, real world application for face recognition is still challenging. In this paper, we propose using the combination of Affine Scale Invariant Feature Transform (SIFT) and Probabilistic Similarity for face recognition under a large viewpoint change. Affine SIFT is an extension of SIFT algorithm to detect affine invariant local descriptors. Affine SIFT generates a series of different viewpoints using affine transformation. In this way, it allows for a viewpoint difference between the gallery face and probe face. However, the human face is not planar as it contains significant 3D depth. Affine SIFT does not work well for significant change in pose. To complement this, we combined it with probabilistic similarity, which gets the log likelihood between the probe and gallery face based on sum of squared difference (SSD) distribution in an offline learning process. Our experiment results show that our framework achieves impressive better recognition accuracy than other algorithms compared on the FERET database.

Keywords

Affine Scale Invariant Feature Transform, Face Recognition, Probabilistic Similarity

1. Introduction

Face recognition has been widely investigated for the last decades, especially for robust face recognition algorithms that are able to deal with real world face recognition, such as identifying individuals from surveillance cameras for public security and annotating people from digital photos automatically. There are some successful commercial face recognition systems available like Google Picasa and Apple iPhoto [1]. However, face recognition research is still far from mature. Earlier face recognition algorithms are only effective under controlled settings, such as the probe and gallery images are frontal. This algorithm fails when it is applied to cases as pose and illumination changes. This paper focuses on the viewpoint invariant face recognition, which identifies a face when probe faces are from different viewpoints, while gallery faces are frontal.

The key issue for face recognition under different viewpoints is the distance between different poses is bigger than the distance between different subjects. One solution is to eliminate the distance between different poses. From among which, face normalization is an effective method to remove the difference in poses. Face normalization can be used as 2D or 3D models. As for the 2D model, Markov random

fields (MRFs) are widely used to find correspondences between the frontal face with the profile probe faces [2,3]. MRFs find 2D displacement by minimizing the energy, which consists of two parts, one is the distance of the corresponding node and the other one represents the smoothness between neighbor nodes. The Lucas-Kanade method is also used for face alignment [4,5]. As for the 3D model, Blanz and Vetter [6] propose an effective 3D morphable method to fit the 3D model to a 2D face, and the fitting shape and texture coefficients were used to recognize faces. The normalization method can be used to construct the frontal face from the probe profile face [2]. It can also be used to directly match between images and the matching score, which represents the similarity between them [3]. These normalization methods are reported as being effective at the cost of a long computation time. It has been reported that two minutes was needed to normalize one face [2]. Marsico et al. [7] proposed a FACE framework to recognize a face for an uncontrolled pose and illumination changes. It detected some keypoints using the SIFT Active Shape Model (STASM) algorithm [8], and constructed a half face by the middle line key points. The rest of the face was reflected from the constructed half face. This easy method was fast, but not robust, because it highly depended on the accuracy of keypoint detection. When the keypoints detection failed, the system performances significantly decreased.

Another solution for face recognition under viewpoint change was to design a new classifier or a new feature. For the new classifier, one-shot similarity (OSS) or two-shot similarity (TSS) are proposed by introducing another dataset, which contains no probe and gallery images [9]. Each dataset contains different images of a single subject or different subjects viewed from a single pose. Similarity scores between two faces are calculated by the model built by one of the faces and the introduced dataset using Linear Discriminant Analysis (LDA) or Support Vector Machine (SVM). Cross-posed face recognition shares a similar concept by introducing a third dataset [10]. Faces from different viewpoints are all linearly represented by the introduced dataset using the subspace method. The similarity between these faces is then calculated indirectly by the linear coefficients. Recently, neural networks [11] and deep learning [12] have been adopted to extract pose invariant features for face recognition. As for new feature extraction, tied factor analysis is proposed to estimate the linear transformation and noise parameters in an 'identity' space [13]. The local descriptor is also an effective way to deal with affine transformation between two images, such as Harris-Affine [14], Hessian-Affine [15], and Affine SIFT [16] algorithms. Recently, the local Gabor binary pattern (LGBP) histogram sequence [17] and high-dimensional local binary pattern (LBP) [18] are proposed as effective local descriptors.

In this paper, we propose using the combination of Affine SIFT and Probabilistic Similarity for face recognition under a large change in viewpoint. In case of a large viewpoint variance, single local descriptors, such as SIFT and LBP, do not work well. The Affine SIFT algorithm tries to detect affine invariant local descriptors with the purpose of increasing the capacity of dealing with a significant pose change. However, the human face is not planar and it contains significant 3D depth. Affine SIFT doesn't work effectively for large pose. To complement this, we combined it with probabilistic similarity, which gets the log likelihood between the probe and gallery face based on SSD distribution in an offline learning process.

The rest of this paper is organized as follows: Section 2 reviews the related work regarding the SIFT algorithm and image alignment. We describe the Affine SIFT algorithm and Probabilistic Similarity for face recognition and their combination ASIFT-PS in Section 3, respectively. Section 4 applies the ASIFT-PS to the FERET database, and presents the experiment results. Finally, we conclude this paper by explaining future work in Section 5.

2. Related Work

2.1 Scale Invariant Feature Transform

Local features are effective methods for conducting matching and recognition as it is robust against occlusion, scale, rotation, or even affine transformation to some extent. Among these algorithms, Scale Invariant Feature Transform (SIFT) is a scale and rotation invariant local feature. It transforms image data into scale-invariant coordinates and localizes the keypoint. Each keypoint is assigned to a descriptor. The major steps for the SIFT algorithm are as described below [19].

1) Scale-space extrema detection: An image is transformed into different scales and size. Extrema are searched by finding maxima and minima over all of scales using a difference of Gaussians scheme, which are invariant to scale and orientation.

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \tag{1}$$

where $D(x, y, \sigma)$ is the difference of Gaussians function, G is the Gaussian function, σ is the scale of Gaussian, and k is the factor of nearby scales. I is the input image. Extrema are detected by comparing a pixel to its 26 neighbors at the current and adjacent two scales (8, 9, and 9 pixels for each scale, respectively) as shown in Fig. 1.

2) Keypoint localization: Extrema are refined by excluding poor localized or low contrast points by checking the refined location, scale, and ratio of principal curvatures. This increases the stability of keypoint localization.
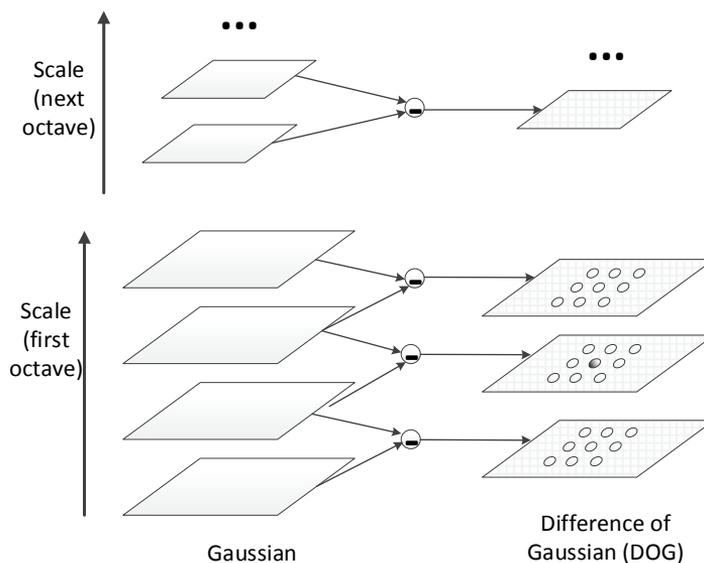

Fig. 1. Illustration of difference of Gaussian (DOG), octaves from bottom to top are generated by down-sampling. The initial image is convolved with Gaussian filter using different scales for each octave. Difference of Gaussian images is generated from this Gaussian filtered image. Extrema are localized by finding the maxima and minima comparing with neighboring pixels in the current scale and adjacent scales as shown on right.

3) Orientation assignment: Each keypoint is assigned to one or more orientations based on a local image gradient histogram. To provide scale and rotation invariance, local image data is transformed to the corresponding orientation and scale for further keypoint descriptor calculations.

4) Keypoint descriptor: The local keypoint descriptor is calculated around each keypoint by the histogram of gradients. The descriptor is transformed into a representation that allows for significant levels of local shape distortion and change in illumination.

A keypoint descriptor is created based on the gradient and orientation in a region around the center keypoint and a Gaussian window weights the region. The region is divided into 4×4 subregions, and the histogram of orientation with eight bins is accumulated for each subregion. Each orientation in the histogram corresponds to the sum of the gradient magnitudes near that direction.

There are several methods that have been proposed, such as best bin first (BBF) [20] and Hough transform [21], for image matching and recognition of the SIFT algorithm. The nearest neighbor is the original and effective matching method for SIFT features. SIFT features are first pre-extracted from gallery images and stored in a database. When matching with a probe image, each SIFT feature from the probe image is compared with all of the gallery features in a database. The nearest neighbor and second nearest neighbor are searched based on the Euclidean distance. The ratio of distances to the second nearest neighbor and nearest neighbor is compared with a threshold. The ratio that is smaller than the threshold is considered to be a matching face.

2.2 Image Alignment

Image alignment is used to find correspondences between two images. Among the algorithms of image alignment, the Lucas-Kanade algorithm is an effective one [4]. This algorithm starts from the non-overlapping division of an image into several subregions. Pixels in the same subregions have the same warp parameters. Let the warp function be $\bar{X} = W(X, P)$ where $\bar{X}$ is the coordinate of pixels of warped subregions with its corresponding pixel coordinate $X = (x, y)$ of the original subregions. $P = [p_1, p_2, \dots p_m]^T$ is the warp parameters with the dimension m. For the affine warp, $m = 6$, and:

$$W(X, P) = \begin{pmatrix} 1 + p_1 & p_3 & p_5 \\ p_2 & 1 + p_4 & p_6 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{2}$$

Let I and T represent the probe image and the gallery image, respectively. Fig. 2 shows these two images captured at two different poses. For each subregion r in T, we try to find a warp that aligns these two images. I_r is the corresponding subregion to T_r after warp transformation. The main objective of alignment is to minimize the error between the T_r and the warped subregions I_r as:

$$E_r = \sum_x \big(I_r\big(W(X, P)\big) - T_r(X)\big)^2 \tag{3}$$

The solution for Eq. (3) is to iterate calculating a ΔP and to update P until P converges. The Lucas-Kanade gives a solution for calculating ΔP by:

$$\Delta P = \mathrm{H}_{img}^{-1} \sum_x \left(\nabla I_r \frac{\partial w}{\partial P}\right)^T (T_r(X) - I_r(W(X, p))) \tag{4}$$

where $\nabla I_r = (\frac{\partial I_r}{\partial x}, \frac{\partial I_r}{\partial y})$ is the gradient of I_r. $\frac{\partial W}{\partial P}$ is the Jacobian of the warp (shown in Eq. (3)). H_{img} is the pseudo Hessian matrix. We can now update the warp parameters $P \leftarrow P + \Delta P$ and iterate till the parameters P converge. This procedure is applied independently for every patch/subregion.

The warp between the probe image and gallery image can be learned online or offline. There are two kinds of online recognition methods. The first one is to calculate a match score for two images based on the warp parameters or alignment errors. Another one is to normalize images by transforming the profile face to its frontal face. Another scheme is off-line alignment. Warp parameters are trained from several sets of images and each set of images are from the same pose.
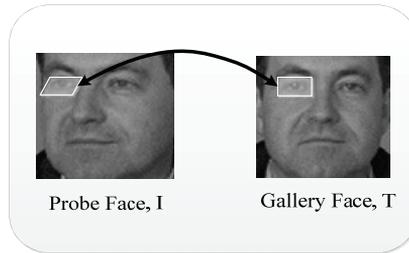

Probe Face, I   Gallery Face, T

Fig. 2. Image to image alignment, an image is divided into several subregions T_r , a warp between two subregions T_r and I_r is calculated by minimizing the alignment error.

3. Proposed ASIFT-PS Algorithm

In this paper, we have proposed using the combination of the Affine SIFT and Probabilistic Similarity for face recognition. Affine SIFT is a stable and effective local descriptors for face recognition. However, the human face contains significant 3D depth. The performance of the Affine SIFT algorithm declines when a significant pose change exists. The Probabilistic Similarity (PS), which is the learning warp transformation between poses considering the 3D depth of a face, can provide complimentary information to the Affine SIFT in significant pose change. However, this alone is not distinctive enough since the warps vary from subject to subject. The combination of these two algorithms increases the capability of the system to recognize a face when there is a significant pose change.

3.1 Affine SIFT

The SIFT is the scale and rotation invariant feature, but it is not an affine invariant. The Affine SIFT is the extension of the SIFT algorithm. There are several parameters for affine transformation, which are:

$$
\begin{aligned}
A &= H_\lambda R_1(\psi) T_t R_2(\phi) \\
&= \lambda \begin{bmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{bmatrix} \begin{bmatrix} t & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}
\end{aligned}
\tag{5}
$$

where λ, R_i and T_t are a scale parameter, rotated angle, and tilted angle, respectively. Fig. 3 shows the geometric relationships of these parameters. The SIFT algorithm is just the scale (λ) and rotation (ψ)

invariant. The other parameters t and ϕ are not invariant. Therefore, the SIFT algorithm is not fully an affine invariant. The Affine SIFT tries to fulfill the t and ϕ invariant, where $t = 1/\cos\theta$.

The Affine SIFT transforms a frontal image into a series of simulated images by the change of longitude ϕ and latitude θ. These simulated images are sampled to achieve a balance between accuracy and sparsity. The ASIFT algorithm is as detailed below [16].

1) The latitude θ is changed to a geometric series $1, a, a^2, \ldots, a^n$, where $a > 1$. According to [16], $a = \sqrt{2}$ is a good compromise between accuracy and sparsity. We also verified this by comparing the face recognition accuracy with various a from $\sqrt{1.5}, \sqrt{2}, \sqrt{2.5}, \sqrt{3}$, and 2. The assignment of $a = \sqrt{2}$ can achieve the best recognition accuracy. For a digital image, the tilt is conducted by a directional t-subsampling with an antialiasing filtered in advance, where $t = 1/\cos\theta$.

2) The longitude ϕ follows an arithmetic series for each tilt, as $0, b/t, \ldots, kb/t$, where $b \approx 72°$ achieves a balance, and k is the last integer satisfying $kb/t < 180°$.
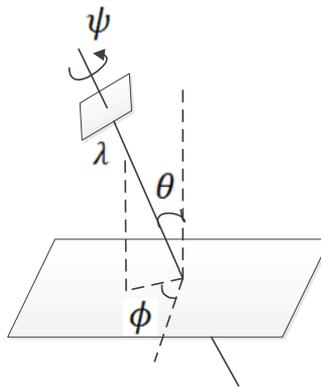

Fig. 3. Geometric relationship of affine decomposition. The λ and ψ are scale and rotation from camera. The θ and ϕ are tilt and rotation of subject, which named latitude and longitude, respectively.

3.2 Probabilistic Similarity

The Affine SIFT improves the face recognition performance to some extent. However, human face is not planar, which has significant 3D depth information. To complement this non-planar feature, we used the probabilistic method as discussed in [5], which is based on image alignment, as described in Section 2.2.

Online warp performs poorly due to its susceptibility to noise, and it takes a long time to converge. The probabilistic method in [5] is an offline method. It learns the distribution of the sum of squared differences (SSDs) between a gallery subregion and its warped probe subregion. SSD is considered to be the similarity between these two subregions. Let's define class $w : w \in \{C, I\}$. Class C means that the gallery and probe images belong to the same subject, while class I means different subjects. The distribution of SSD for each class can be approximated as:

$$p(s_r \mid \phi, w) = \frac{1}{\gamma_r^w \sqrt{2\pi} s_r} \exp\left(-\frac{1}{2}\left(\frac{\log(s_r) - \upsilon_r^w}{\gamma_r^w}\right)^2\right)$$

(6)

where s_r is the SSD of subregions r between the gallery image and probe images. υ_r^w and γ_r^w are the log-means and the log-standard-deviations of the SSD for class w ($w \in \{C, I\}$), respectively. These two parameters can be learned offline from the SSD histogram using some subset of the database.

After the offline learning, we can get the distribution of SSD for each subregion. For a new probe face, the log-likelihood ratio L of the probe image and gallery image belonging to the same subject is:

$$L = \sum_r (\log p(s_r \mid w = C, \phi_p) - \log p(s_r \mid w = I, \phi_p)) \tag{7}$$

After calculating probability L with each gallery, we can get the similarities of the probe image with each gallery image.

3.3 ASIFT-PS Framework

In our framework, the Affine SIFT is first adopted to each gallery face (frontal face), and the probe faces use the SIFT algorithm. The Affine SIFT is used to detect keypoints and local features for the gallery face, and are stored in the keypoints database. For the recognition part, the SIFT algorithm is used to compute keypoints and local features for each probe face. This SIFT keypoints are compared with the Affine SIFT keypoints that are stored in the database. Let N be the total number of subjects to be recognized in the gallery. Each probe face is compared with these N subjects, resulting in a vector $\{M_1, \dots, M_i, \dots M_N\}$, where M_i is the number of matching keypoints with subject i in the gallery.

In order to combine the PS algorithm, we need to first learn the υ_r^w (log-means) and γ_r^w (log-standard-deviations) of the SSD for cases that two faces belong to the same subject and not, respectively. In our experiment, we used half of the face images to train these parameters. Let N be the number of half training subjects. For each subregion r between a pair of poses, we have N pairs of the same subject and $N^*(N-1)$ pairs of different subjects. The SSD of these pairs are calculated separately from the same subject and different subjects. υ_r^C and γ_r^C are then obtained as the log-means and log-standard-deviations of the SSD of pairs with the same subjects. While υ_r^I and γ_r^I are the log-means and log-standard-deviations of the SSD of pairs with different subjects. Based on the $\upsilon_r^C, \gamma_r^C, \upsilon_r^I$, and γ_r^I we can calculate the probability similarity between a probe face with a gallery face using Eqs. (6) and (7). As a result, for N subjects to be recognized, a vector $\{L_1, \dots, L_i, \dots L_N\}$ is obtained for each probe face, where L_i is the probability similarity with subject i in the gallery.

The final similarity between a probe face and a gallery face i is calculated as:

$$S_i = M_i + \lambda_i, (\lambda > 0) \tag{8}$$

where M_i and L_i are the number of matching and probabilistic similarities with gallery image i. The λ is used to balance the value range between the matching number and probability. In our experiment, we varied λ from 1 to 15 with an interval of 1, and evaluated the face recognition rate. Fig. 4 shows the results with various λ and illustrates that $\lambda = 10$ can achieve a good performance.

We can summarize our algorithm as follows:

Algorithm: Face Recognition based on ASIFT-PS

Offline process:

(1) Let *N* be the number of training subjects, where we have N pairs of the same subject and N*(N-1) pairs of different subjects, calculating the υ_r^C, γ_r^C, υ_r^I and γ_r^I for each sub-region *r* based on the SSD between these pairs of images.

(2) Generate several affine transformed faces for each gallery face by varying the longitude ϕ and latitude θ, using SIFT to detect the keypoints of these simulated faces and store K_g^i for gallery face *i*.

Online process:

For each probe face:

    Calculate its keypoint K_p using the SIFT algorithm.

    For each gallery face *i*:

        a) Match keypoint K_p with keypoints K_g^i and get the number of matching keypoints M_i

        b) Align the probe face and gallery face *i*, and get the SSD of each subregion. Based on υ_r^C, γ_r^C, υ_r^I, and γ_r^I, which are calculated in the offline process, estimate the probability similarity L_i using Eqs. (6) and (7);

        c) Calculate the similarity S_i between the probe face and gallery face *i* using Eq. (8).

    End

    Find the maximum S_i and the index of recognize face $r = arc\max_i S_i$.
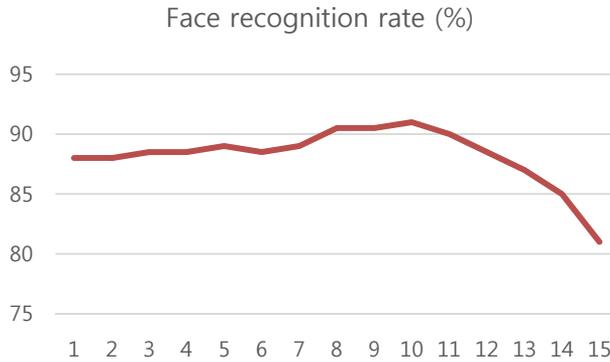
End

Fig. 4. Face recognition accuracy with various λ from 1 to 15 with a step 1.

4. Experimental Results

In our experiments, we used the FERET [22] grey database to evaluate our algorithm. This database contains 200 subjects and each subject contains nine images captured from different poses. For each subject, we used the frontal image as a gallery, and the other eight pose images as probe images, the pose angles of which are 60°, 40°, 25°, 15°, -15°, -25°, -40°, and -60°, respectively. These pose angles correspond to the horizontal viewpoint change from left to right. Fig. 5 shows these face images in the FERET database.

The proposed ASIFT-PS algorithm is compared with the LBP [23], LGBP [17], OSS [9], SIFT [19], Probabilistic Similarity [5], and ASIFT [16]. The parameters used in our experiment for the SIFT algorithm are as follows: the image is resized to a resolution of 400×400 and the ratio for the nearest neighbor is set to 0.8. For Probabilistic Similarity, the image is divided into 20×20 non-overlapping subregions for alignment. We used a CPU is Intel Core i7-4790 3.6 GHz, RAM is 8 GB to perform all of the experiments. The system that we used is a 64-bit Windows 7 Enterprise SP1.

Fig. 5. Face images in FERET database with varying pose from (a) 0°, (b) 60°, (c) 40°, (d) 25°, (e) 15°, (f) -15°, (g) -25°, (h) -40°, and (i) -60°, respectively.

Table 1. Comparison of recognition accuracy with SIFT, PS, ASIFT on FERET database

Algorithm	-40°	-25°	-15°	15°	25°	40°	Average	Time (ms)
LBP	28	64	90	86	57	42	61	12
LGBP	47	71	94	94	66	49	70	786
OSS	37	70	92	87	73	44	67	752
SIFT	62	94	99.5	99.5	94.5	52	84	695
PS	48	57.5	73.5	82.5	76.5	48	64	18
ASIFT	63	95	99.5	99.5	95.5	66.5	87	1156
ASIFT-PS	74	96.5	99.5	99.5	98	78.5	91	1283

SIFT=Scale Invariant Feature Transform, PS=Probabilistic Similarity, ASIFT=Affine-SIFT, LBP=local binary pattern, LGBP=local Gabor binary pattern, OSS=one-shot similarity.

We modified the original Affine SIFT so as not to spend computational time on invalid cases. We did this because the Affine SIFT generates a lot of viewpoints from a single frontal image. When setting the number of tilt to seven, the Affine SIFT generates 61 viewpoints. It takes too much time to match the keypoints with all of these viewpoints, and some of them are not that significant for face recognition because the face is detected only around special angles. Therefore, we restricted the viewpoints in a discrete series with only horizontal and vertical directions. This restriction would not reduce the performance much because the SIFT is effective when the pose difference is within 30 degrees. In our experiments, three and five viewpoints were examined in the horizontal and in the vertical direction.

This results in a total of 15 viewpoints generated.

Table 1 shows the face recognition rate and computational time of face recognition with LBP, LGBP, OSS, ASIFT, SIFT, and PS on the FERET database. The computational time was calculated by averaging the matching time of each test face (Matching time is the time it takes to calculate one similarity). From among the compared algorithms, the ASIFT-PS achieved the best recognition performance and its computational time is the longest as well. However, it is worth noting that the computational time of ASIFT-PS does not have much of a difference with ASIFT, while the performance of ASIFT-PS is much better. We also noticed that SIFT could get similar results with ASIFT-PS when a pose degree was between -15° to 15°. However, ASIFT-PS achieved better results than SIFT or PS under a significant pose difference. The performance of the PS algorithm alone is not good. However, when combined with the ASIFT, the performance significantly increased. This is because the likelihood probability of the PS algorithm in a significant pose complement the performance deteriorates of ASIFT in large pose change. The ASIFT-PS algorithm didn't increase the computational time by much because PS is performed in an offline manner. The face recognition rate was plotted for the FERET database in Fig. 6. The performance of PS was the worst. However, the ASIFT-PS, which is a combination of ASIFT and PS, showed the highest performance. It achieved about 10% in gains of recognition accuracy in the average under significant pose degrees.
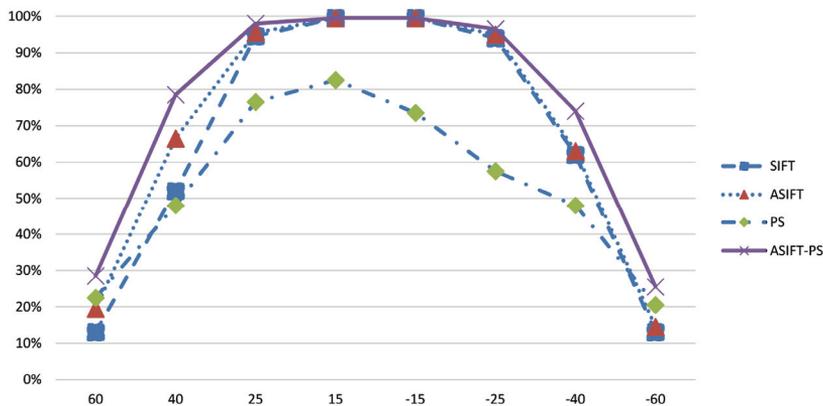

Fig. 6. Recognition rate of face with degrees varying from 60°, 40°, 25°, 15°, -15°, -25°, -40°, -60° in FERET database. Algorithms are Scale Invariant Feature Transform (SIFT), Probabilistic Similarity (PS), Affine-SIFT (ASIFT) and proposed ASIFT-PS.

5. Conclusions

In this paper, the ASIFT-PS was used for face recognition when a viewpoint change occurs. The Affine SIFT is an extension of the SIFT algorithm. The SIFT algorithm is the scale and rotation invariant, which is powerful for small viewpoint changes in face recognition. In our scheme, the Affine SIFT was only used for gallery faces, which generated a series of different viewpoints using affine transformation. In this way, it enabled viewpoint difference between gallery faces and probe faces. However, a human face is not planar and contains significant 3D depth. The Affine SIFT does not work well for a significant pose difference. The strength of the proposed method is that by combining the Affine SIFT and Probabilistic Similarity, the

limitations of the Affine SIFT is eliminated and the Probabilistic Similarity's statistical information compensates for the deficiencies of the SIFT algorithm. Thus, the accuracy of facial recognition increases by more than 10%, compared to other algorithms. To complement this, we combined the Affine SIFT with probabilistic similarity, which obtained the log likelihood between the probe and gallery face based on SSD distribution in an offline learning process. The FERET database was used to test the proposed ASIFT-PS. The results of our experiment showed that SIFT could achieve comparable results with the Affine SIFT when a pose degree was between -15° to 15°, but the ASIFT-PS achieved about 10% better on average as compared to SIFT under a large pose different.

Acknowledgement

References

[1] G. Hua, M. H. Yang, E. Learned-Miller, Y. Ma, M. Turk, D. J. Kriegman, and T. S. Huang, "Introduction to the special section on real-world face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 1, pp. 1921-1924, 2011.

[2] H. T. Ho and R. Chellappa, "Pose-invariant face recognition using Markov random fields," *IEEE Transactions on Image Processing*, vol. 22, no. 4, pp. 1573-1584, 2013.

[3] S. R. Arashloo and J. Kittler, "Energy normalization for pose-invariant face recognition based on MRF model image matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 6, pp. 1274-1280, 2011.

[4] S. Baker and I. Matthews, "Lucas-Kanade 20 years on: a unifying framework," *International Journal of Computer Vision*, vol. 56, no. 3, pp. 221-255, 2004.

[5] A. B. Ashraf, S. Lucey, and T. Chen, "Learning patch correspondences for improved viewpoints invariant face recognition," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR2008)*, Anchorage, AK, 2008, pp. 1-8.

[6] V. Blanz and T. Vetter, "Face recognition based on fitting a 3D morphable model," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1063-1074, 2003.

[7] M. De Marsico, M. Nappi, D. Riccio, and H. Wechsler, "Robust face recognition for uncontrolled pose and illumination changes," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 43, no. 1, pp. 149-162, 2013.

[8] S. Milborrow and F. Nicolls, "Locating facial features with an extended active shape model," in *Proceedings of 10th European Conference on Computer Vision*, Marseille, France, 2008, pp. 504–513, 2008.

[9] L. Wolf, T. Hassner, and Y. Taigman, "Effective unconstrained face recognition by combining multiple descriptors and learned background statistics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 10, pp. 1978-1990, 2011.

[10] A. Li, S. Shan, and W. Gao, "Coupled bias–variance tradeoff for cross-pose face recognition," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 305-315, 2012.

[11] Y. Zhang, M. Shao, E. K. Wong, and Y. Fu, "Random faces guided sparse many-to-one encoder for pose-invariant face recognition," in *Proceedings of IEEE International Conference on Computer Vision (ICCV)*, Sydney, Australia, 2013, pp. 2416-2423.

[12] Z. Zhu, P. Luo, X. Wang, and X. Tang, "Deep learning identity preserving face space," in *Proceedings of IEEE International Conference on Computer Vision (ICCV)*, Sydney, Australia, 2013, pp. 113-120.

[13] S. J. D. Prince, J. Warrell, J. H. Elder, and F. M. Felisberti, "Tied factor analysis for face recognition across large pose differences," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 6, pp. 970-984, 2008.

[14] K. Mikolajczyk and C. Schmid, "Scale and affine invariant interest point detectors," *International Journal of Computer Vision*, vol. 60, no. 1, pp. 63-86, 2004.

[15] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. V. Gool, "A comparison of affine region detectors," *International Journal of Computer Vision*, vol. 65, no. 1-2, pp. 43-72, 2005.

[16] J. M. Morel and G. Yu, "ASIFT: a new framework for fully affine invariant image comparison," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 438-469, 2009.

[17] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang, "Local Gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition," in *Proceedings of IEEE International Conference on Computer Vision (ICCV)*, Beijing, China, 2005, pp. 786-791.

[18] D. Chen, X. Cao, F. Wen, and J. Sun, "Blessing of dimensionality: high-dimensional feature and its efficient compression for face verification," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Portland, OR, 2013, pp. 3025-3032.

[19] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.

[20] J. Beis and D. G. Lowe, "Shape indexing using approximate nearest-neighbour search in high-dimensional spaces," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Juan, Puerto Rico, 1997, pp. 1000-1006.

[21] P. V. Hough, "Method and means for recognizing complex patterns," U.S. Patent 3069654, 1962.

[22] P. J. Phillips, H. Wechsler, J. Huang, and P. Rauss, "The FERET database and evaluation procedure for face-recognition algorithms," *Image and Vision Computing*, vol. 16, no. 5, pp. 295–306, 1998.

[23] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: application to face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037-2041, 2006.

Yongbin Gao

He received B.S. and M.S. degrees in School of Information Technology from Jiangxi University of Finance and Economics, China in 2010 and 2013, respectively. Since September 2013, he is with the Division of Computer Science and Engineering from Chonbuk National University as a Ph.D. candidate.

Hyo Jong Lee http://orcid.org/0000-0003-2581-5268

He received B.S., M.S., and Ph.D. degree in Computer Science from University of Utah, U. S. in 1986, 1988, and 1991, respectively. Since 1991, he is with the Division of Computer Science and Engineering from Chonbuk National University as a professor. His current research interests include image processing, computer vision, medical imaging, and parallel processing.