Optimal Pose for Face Recognition

Xiaoming Liu[1] Tsuhan Chen[2] Jens Rittscher[1]

[1]Visualization and Computer Vision Lab
General Electric Global Research Center
Schenectady, NY, 12309
{liux,rittsche}@research.ge.com

[2]Advanced Multimedia Processing Lab
Carnegie Mellon University
Pittsburgh, PA, 15213
tsuhan@cmu.edu

Abstract

Researchers in psychology have well studied the impact of the pose of a face as perceived by humans, and concluded that the so-called 3/4 view, halfway between the front view and the profile view, is the easiest for face recognition by humans. For face recognition by machines, while much work has been done to create recognition algorithms that are robust to pose variation, little has been done in finding the most representative pose for recognition. In this paper, we use a number of algorithms to evaluate face recognition performance when various poses are used for training. The result, similar to findings in psychology that the 3/4 view is the best, is also justified by the discrimination power of different regions on the face, computed from both the appearance and the geometry of these regions. We believe our study is both scientifically interesting and practically beneficial for many applications.

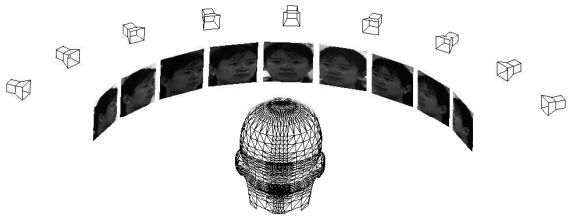

Figure 1. Problem setup: Given multiple poses of a human face, which is the *optimal training pose* **that will lead to the best pose-robust recognition performance? Why is it optimal?**

1 Introduction

For decades face recognition has been widely studied in the computer vision community. Comprehensive surveys of machine recognition techniques can be found in [22]. Face recognition research is primarily concerned with handling *variations*, such as pose, illumination and expression. Pose variation is the hardest to model and therefore contributes to most of the recognition errors [16]. Thus, pose-robust face recognition is receiving more attention recently [2, 6, 11]. However, although many approaches have been proposed to achieve better performance in terms of pose-robust recognition, little work has been done in understanding the intrinsic capabilities of different poses in the context of face recognition by machines.

The hypothesis of this paper is that the overall capability of recognizing a face in an arbitrary pose depends on the training pose being chosen, assuming only one pose is used for training. That is, as shown in Figure 1, among various poses of the same face, using one of them as the training pose would lead to better pose-robust recognition performance than using other poses. This paper suggests approaches to test this hypothesis, find the optimal training pose, and provide justification regarding the finding.

Our work is motivated by the research in psychology, where it is often reported that faces are better recognized by humans if learned or tested in the 3/4 view, halfway between the front view and the profile view [12]. Obviously this topic is composed of two relevant questions. One is how the use of different poses in learning (or training) affects the test performance, which is called the *Optimal Representative Pose Problem (ORPP)* in this paper. The other is how the use of different poses as the test data affects the performance, which is called the *Optimal Recognizable*

Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)
0-7695-2597-0/06 $20.00 © 2006 IEEE

Pose Problem (OCPP). In the vision community, it is generally accepted that when the training pose has been chosen, the amount of angular rotation between the training and test pose will be the primary factor affecting recognition performance. In addition, the question of recognition in the *same* pose between the training and test is not actually a pose-dependent problem, since under the same pose other types of variations, such as expression or lighting, will affect the recognition instead of the pose. For these reasons, our paper will focus on ORPP. That is, compared to psychological findings, what conclusion can we draw on ORPP when the face is recognized by machines? As a scientific problem, it is interesting to see whether the conclusion in computer vision validates or contradicts that of psychology. Similar to almost all psychology studies [3, 7, 10, 12], only the horizontally rotated poses are studied in this paper.

This study will benefit practical applications in that it offers theoretical support on why faces should be captured at a certain pose in practice. For example, in law enforcement applications, mug shots are generally taken at either front or profile view. Given the fact that mug shots will be used in face recognition applications, they should preferentially be taken at the optimal pose, i.e., the 3/4 view, as concluded in both this paper and prior psychology study.

The paper is organized as follows. Section 2 reviews literature in both psychology and computer vision. In Section 3, an experimental approach is used to find the optimal training pose for face recognition. In Section 4, we provide justification on why a certain pose has advantages over others in recognition. Conclusions are given in Section 5.

2 Previous work

2.1 Psychology

Researchers [7, 12] in psychology have been working on the role of pose in human face recognition for decades. It is well established that certain views of objects are easier to remember and are usually referred as *canonical views* [15]. For human faces, a 3/4 view is often an obvious candidate for a canonical view. A number of researchers conclude that using the 3/4 view at learning leads to superior levels of recognition [12]. However, there is a lack of consistency about the advantage of the 3/4 view in the literature [10], since some studies conclude that there is no noticeable advantage of using the 3/4 view as the learning pose [3].

As Liu and Chaudhri [10] have pointed out, the reason for the inconsistency might come from the experimental methodology undertaken. When a face in the 3/4 view generalizes better to a profile view than a front view, the effect is likely contributed by the *smaller* amount of angular rotation for the 3/4 view. Because with its unique halfway position between the front and profile view and given the symmetry of the face, the 3/4 view has, on average, the smallest amount of angular rotation to all other views. For this reason, the effect of angular rotation should be seen more as an extrinsic factor for the 3/4 view advantage. If the recognition experiments are conducted using all possible poses, the angular rotation will contribute to the conclusion. Hence another experimental setup is to exclude such extrinsic factor and focus on potential advantage due to the intrinsic properties of the 3/4 view, by testing poses only within the range of certain angular rotation with respect to the training pose, instead of all possible poses. In this paper, we will conduct experiments using both methodologies.

2.2 Computer vision

As shown in the recent survey [22], most of the vision literature in face recognition focuses on proposing new algorithms. However, relatively little work has been done in understanding the relationship between variation and recognition performance. For example, one type of approach is to treat the face image under a certain pose as one sample in a high-dimensional space, and learn the relation between a training pose image and test pose images by building a mapping function between them. Given a test image with an arbitrary pose, a recognition-by-synthesis approach is applied. That is, we can either map this test image into the training view [8], or map each of the training images into the same pose as the test image [13], based on the learned mapping function. However, to our knowledge, few studies have been performed regarding the level of confidence when each pose is mapped into other poses.

Yacoob and Davis [21] are among the first to investigate how expression variation affects recognition. They conclude that smiling faces should be preferred if a system has a choice in the selection of faces to use in training and recognition. For pose variation, Graham and Allinson [5] have done interesting work in analyzing the recognition performance under different pairs of training and test poses using their eigensignatures method. They find that the 40° view seems to have the best performance when it is used in training or test. However, no explanation or other measurement has been provided regarding the results. Lee and Sohn [9] represent the level of the influence of a given view over nearby views by using a "quasi-view" size, which is the size of the projected face model. However, there is no evidence suggest that the size of the projected face is a dominant factor for the recognition capability. Weber *et al.* [19] show that in their viewpoint-invariant face detection approach, the best detection performance is achieved by training at 30° viewpoint intervals. However, no explicit analysis is performed regarding this finding. Weinshall and Werman [20] analytically study the view likelihood and stability for general objects, and conclude that the most likely and stable

view is the flattest view of the 3D shape. Their finding is consistent with our geometry analysis and our analysis is specific to faces.

3 Experimental approach

As stated in Section 1, we hypothesize that some poses are more representative than others. Hence if they are used as the training poses, the recognition performance will be better. To test this hypothesis, we can conduct recognition experiments by using a particular face recognition algorithm. Given a face database with ground truth on pose, we can exhaustively choose each pose as the training data, and evaluate its corresponding recognition performance using other poses as the test data. Although this approach is easy to perform, one concern is that the conclusion drawn from this approach might depend on the specific face recognition algorithm being chosen. In order to mitigate this concern, we choose algorithms from different types of recognition methods.

Among many existing algorithms for face recognition, some are generic algorithms and do not take special attention to specific variation, such as pose. Eigenface [18] and Fisherface [1] are good examples of this type of algorithms. On the other hand, there is another family of algorithms which target pose-robust face recognition [11, 13]. We call them *pose-aware algorithms* since they are aware of pose variation and explicitly compensate for it using specific methods.

In our experimental approach, eigenface is chosen as an example of the generic algorithms, while the geometry assisted probabilistic mosaic algorithm (GAPMA) [11] is used as an example of the pose-aware algorithms.

3.1 Face database

We use the CMU PIE database [17] in our study, since it has a relatively dense sampling of poses, and has labeling of the ground truth for each pose. The PIE database consists of face images of 68 subjects under different pairs of poses and illuminations. In this paper, we use the pose subset of this database, which contains 9 pose images for each of 68 subjects captured under the neutral illumination conditions. Sample images from one subject are shown in Figure 2. According to the cameras' 3D location included in the meta data, the true pose of these images span approximately -64° to 64° horizontally and differ about 16° between neighboring poses.

3.2 Eigenface algorithm

Detailed information about the eigenface algorithm can be found in [18]. We use a PIE subset of 34 subjects each with 9 poses as the experimental dataset. For each image in the dataset, we manually crop the face area and normalize it to 64 by 64 pixels. In order to remove the influence of the background, we also apply a mask to all images such that the 4 triangular-shaped corners are filled with zeros.

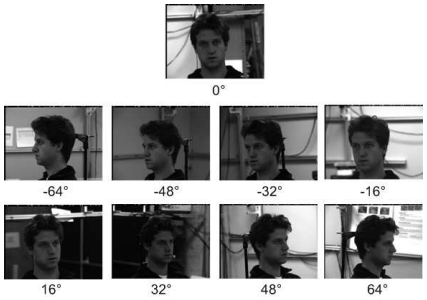

Figure 2. Sample images of one subject in the PIE database.

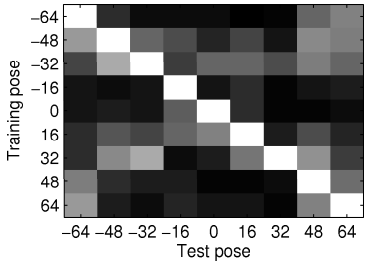

Figure 3. Recognition rate matrix of the eigenface algorithm for all pairs of training and test poses. Each element has a value between 0% and 100%. The diagonal elements are always 100% since the same training image is used for test.

We iteratively use each of the 9 poses of all subjects as the training data to generate an eigenspace. The remaining 8 poses from all subjects are then used as the test data. The nearest neighbor classifier is applied during the testing stage. Since there are 34 training images (34 subjects each with 1 pose image) in total, it is possible to obtain an eigenspace whose number of eigenvectors varies from 1 to 33. We test all different numbers of eigenvectors, and report the one with the best recognition performance. By assuming the symmetry of the face, the test face is flipped horizontally whenever its flipped pose has less angular deviation from the training pose. The recognition rates for all pairs of training and test poses are computed as shown in Figure 3.

Since we are interested in how recognition performance

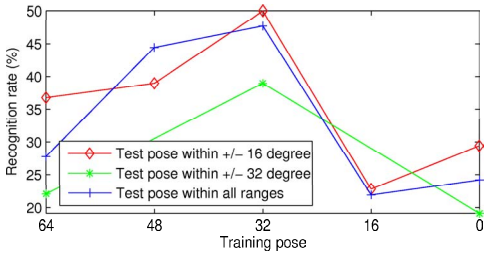

Figure 4. Average recognition rate of the eigenface algorithm with respect to different training poses. Each point is computed via an equation similar to (3). In order to exclude the factor of angular rotation, the test poses are chosen to be within the range of left & right 16° or 32° of the training pose. The factor of angular rotation contributes to the results when all test poses are evaluated. In both cases, the 32° training pose provides the best recognition performance.

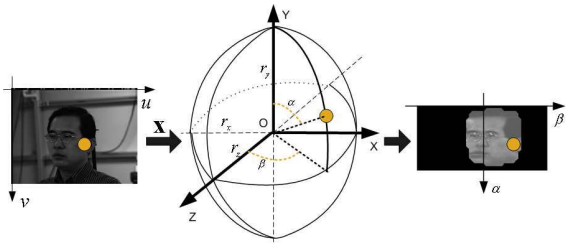

Figure 5. A face image is mapped onto the surface of a 3D ellipsoid according to the estimated mapping parameter X. This mapping compensates for pose variation.

changes with different training poses, one simple way to evaluate performance is to average each row of the recognition rate matrix. However, as we discussed above, there is one experimental methodology excluding the factor of angular rotation by only testing the pose that is within a certain angular rotation of the training pose. Applying this methodology is equivalent to horizontally averaging the elements within a fixed-width strip whose center is the diagonal element. Eventually based on the symmetric assumption, the recognition results of opposite poses are averaged. For example, the results of 32° and -32° are averaged and are treated as the final result of 32°. Results from both methodologies are shown in Figure 4. Notice that when the averaging range is 32°, the averaging at 16° and 48° training poses are not computed because part of their 4 neighboring poses within the 32° range do not truly span the angular rotation as 0°, 32°, and 64° do. From Figure 4, it is obvious that the 32° pose leads to better performance than other poses.

3.3 Geometry assisted probabilistic mosaic algorithm

The basic idea of GAPMA is to use geometry to compensate for pose variation. As shown in Figure 5, the human head is approximated with a 3D ellipsoid model. Thus, any face image is a 2D projection of such a 3D ellipsoid at a certain pose. In this approach, both training and test images are back-projected to the surface of the 3D ellipsoid, according to their estimated poses, to form the texture maps. Thus the recognition can be conducted by comparing the texture maps instead of the original images, as is done in traditional face recognition. Furthermore, the texture map is represented as an array of local patches, which allows the training of a probabilistic model for comparing corresponding patches. Details of this approach can be found in [11].

Based on the same testing scheme as the eigenface algorithm, we conduct experiments on the PIE database using GAPMA. A subset of 34 subjects each with 9 poses are used for experiments. The remaining subjects are used for training the probabilistic model. Similarly, we iteratively use each of the 9 poses of all 34 subjects as the training data and the rest as the test data. The recognition rate matrix for all pairs of training and test poses is shown in Figure 6. Also, the average recognition rate for each training pose is shown in Figure 7, computed in the same way as Figure 4.

Together with the results from the eigenface algorithm, a number of observations can be made. First, for both algorithms, the use of 32° as the training pose results in the best performance in the recognition stage, including or excluding the factor of angular rotation. Second, the advantage of 32° seems more obvious for generic face recognition algorithms, while not dramatic for pose-aware algorithms, especially when excluding the factor of the angular rotation. This is natural since the overall influence of pose variation on face recognition is reduced because the pose-aware algorithms explicitly compensate for pose variation.

4 Why is 32° the optimal pose?

We have shown that for both generic and pose-aware algorithms, 32° is the optimal training pose that leads to the best recognition performance. However, one might want to ask *why* 32° is optimal. Researchers in psychology [10] have worked on this subject for decades and there is still no conclusive and convincing explanation on why a certain pose has advantages over others. O'Toole *et al.* [14] conclude that both 3D geometry and 2D surface appearance

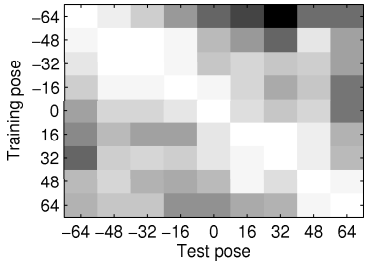

Figure 6. Recognition rate matrix of GAPMA. The performance is much better than that of the eigenface (Figure 3).

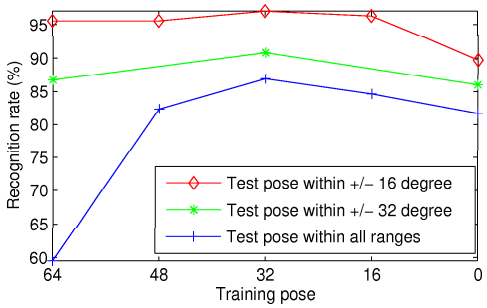

Figure 7. Average recognition rate of GAPMA. The 32° training pose provides the best recognition performance, especially when testing on all possible poses.

contribute to human face recognition across poses. We attempt to justify the optimal pose from the computer vision point of view, where appearance and geometry are also the two fundamental factors for object recognition. Our justification is based on these two factors as well. Essentially, given each training pose, we will compute its discrimination power against test poses based on the facial appearance and geometry.

4.1 Facial appearance evidence

How important is the facial appearance in discriminating facial identity? The basic assumption is that different parts/patches of the face contribute differently to face recognition. If we could quantify such differences in terms of discrimination power for all pairs of training and test poses, we can find out which pose is more powerful in recognizing other poses. The Fisher ratio [4] is computed since it is a conventional way of measuring the discrimination power [11].

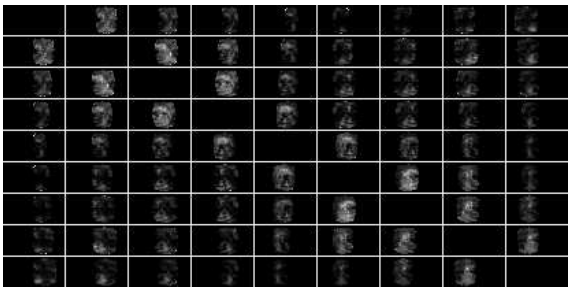

Figure 8. Fisher ratio maps (FRM) for all pairs of training and test poses. Each pixel, $f_{i,j}^{m,n}$, is the Fisher ratio of the patch (i,j) between the training pose ϕ_m and test pose ϕ_n. Brighter pixels indicate the higher Fisher ratio and greater contribution to recognition. Vertical axis corresponds to the training poses from -64° (top) to 64° (bottom). Horizontal axis corresponds to test poses from -64° (left) to 64° (right). Notice the relatively brighter maps around the -32° (3^{rd}) row and the 32° (7^{th}) row.

Given the PIE database of L $(L=68)$ subjects each with 9 poses, let us introduce how to compute Fisher ratio for each pair of training and test poses. After performing geometry mapping for each image $\mathbf{f}(l,\phi_m)$, where l, ϕ_m are the index of subjects and poses respectively, the resultant texture map is represented as an array of local patches $\mathbf{s}_{i,j}(l,\phi_m)$. Fisher ratio will be computed based on the texture maps, instead of the original face images.

Suppose we treat one of the 9 poses, ϕ_m, as the training pose, we need to study how the distance measure of corresponding patches between ϕ_m and all other 8 test poses changes. This is done by fixing one patch (i,j) and one particular test pose ϕ_n, and calculating the mean square error (MSE) of one patch between all subjects in the pose ϕ_m and ϕ_n. MSE is used here since it is a common way of computing distance measurement in appearance-based object recognition. A distance matrix is generated where each element indicates the MSE of the same patch (i,j) between one pair of subjects. The distribution of the diagonal elements of this distance matrix is an indication of the intra-subject variations, while that of the off-diagonal elements is an indication of the inter-subject variations. We explicitly model these two distributions as Gaussian distributions, whose means are computed as:

$$\mu_{i,j}^{m,n^{\text{same}}} = \frac{1}{L}\sum_{l} \|(\mathbf{s}_{i,j}(l,\phi_m) - \mathbf{s}_{i,j}(l,\phi_n)\|^2$$

$$\mu_{i,j}^{m,n^{\text{diff}}} = \frac{1}{L(L-1)}\sum_{l\neq k} \|(\mathbf{s}_{i,j}(l,\phi_m) - \mathbf{s}_{i,j}(k,\phi_n)\|^2 \qquad (1)$$

where $\mu_{i,j}^{m,n^{\text{same}}}$ and $\mu_{i,j}^{m,n^{\text{diff}}}$ are the mean of intra-subject and inter-subject variations for the patch (i,j) between the training pose ϕ_m and the test pose ϕ_n. In addition, the standard deviation $\sigma_{i,j}^{m,n^{\text{same}}}$ and $\sigma_{i,j}^{m,n^{\text{diff}}}$ of these two distributions are computed as well. Given two Gaussian distributions, the Fisher ratio [4] is defined as:

$$f_{i,j}^{m,n} = \frac{(\mu_{i,j}^{m,n^{\text{diff}}} - \mu_{i,j}^{m,n^{\text{same}}})^2}{\sigma_{i,j}^{m,n^{\text{same}}\,2} + \sigma_{i,j}^{m,n^{\text{diff}}\,2}}$$

Note that this Fisher ratio is computed based on one fixed patch (i,j) and one pair of training and test pose, ϕ_m and ϕ_n. By varying the location of the patch across all facial area, a *Fisher ratio map (FRM)* is obtained between one pair of poses. Finally for all pairs of training and test poses, an array of FRM is generated as shown in Figure 8.

In order to have a quantitative evaluation for the Fisher ratio of each training pose, we compute the average of Fisher ratio $f^{m,n}$ over non-empty patches, which are visible from both training and test views, for each map as in (2). Then the Fisher ratio for each training pose is averaged across the test poses, with different ranges r as in (3). We show the resultant $h^{m,r}$ for $m = 1,2,...,5$ and $r = 1,2,8$ in Figure 9. It is apparent that when excluding the factor of angular rotation, i.e., the test pose is within the range of 16° or 32° with respect to the training pose, 32° has higher Fisher ratio than other training poses. Since higher Fisher ratio implies better recognition performance, this result is consistent with our empirical conclusion in Section 3. However, when all test poses are considered, all three poses around 32° seem to work equally well.

$$f^{m,n} = \frac{1}{\sum_{i,j}(f_{i,j}^{m,n} > 0)} \sum_{i,j} f_{i,j}^{m,n} \quad (2)$$

$$h^{m,r} = \frac{1}{2}\left(\frac{1}{2r} \sum_{n}^{0<|m-n|\le r} f^{m,n} + \frac{1}{2r} \sum_{n}^{0<|10-m-n|\le r} f^{10-m,n}\right) \quad (3)$$

4.2 Geometric evidence

The human face has a unique 3D geometry. If the facial geometry were a perfect sphere, the optimal training pose would be *only* dependent on the facial appearance. Our hypothesis is that because of 3D facial geometry, the 2D face images projected from different poses will have different capabilities to infer the face at other poses. Thus, we will study the inference capability between each pair of training and test poses based on the 3D geometry. Again this inference capability will depend on the location of the facial patch, and eventually an *inference capability map (ICM)* similar to FRM in Figure 8 will be obtained, which provides a natural way of integrating the appearance and geometry information.

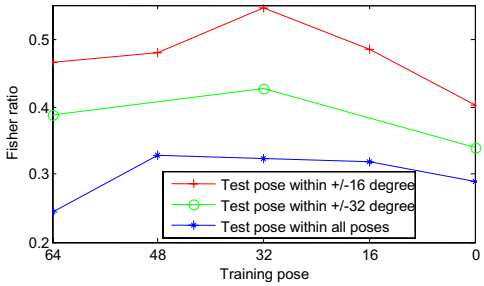

Figure 9. Average Fisher ratio map (FRM) with respect to different training poses. The 32° view advantage can be seen when excluding the factor of angular rotation.

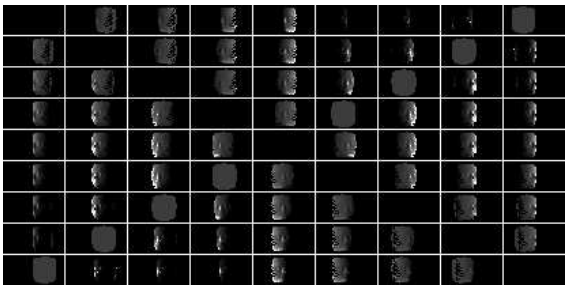

Figure 10. Inference capability maps (ICM) for all pairs of training and test poses. Each pixel, $g_{i,j}^{m,n}$, is computed via (4).

Our study is based on a generic 3D face mesh model $\mathbf{M} = \{\mathbf{t}_u\}$, which is composed of a set of triangles $\mathbf{t}_u$. Each triangle is described by three vertices in 3D. The basic approach is as follows. First, we build the correspondence between the set of triangles and the array of patches in the appearance model. This can be achieved by computing the horizontal and vertical angles of the vector connecting the center of the 3D face model and the center of the triangle. Then the angles can be easily converted into the horizontal and vertical index of the corresponding patch. The result of this step is a mapping function $f(u) = [i,j]$, which maps the triangle index u into the index of the corresponding patch $[i,j]$.

Second, we project the 3D geometry model into the 9 different views used in the PIE database, by assuming an orthographic camera model. Hence another set of triangles $\{\mathbf{t}_u^m\}$ are obtained for each pose ϕ_m on the image plane. Now given the training pose ϕ_m, and the test pose ϕ_n, we can find

the matched triangles between $\{\mathbf{t}_u^m\}$ and $\{\mathbf{t}_u^n\}$. One way to measure the inference capability is to measure the ratio of the total area of the matched triangles between two poses that correspond to the same appearance patch, as follows:

$$g_{i,j}^{m,n} = \frac{\sum_u^{f(u)=[i,j]} \Delta(\mathbf{t}_u^m)}{\sum_u^{f(u)=[i,j]} \Delta(\mathbf{t}_u^n)} \quad (4)$$

where the numerator and the denominator are the total area of projected triangles that correspond to the appearance patch with index $[i, j]$, at the training pose ϕ_m and the test pose ϕ_n respectively. It can be seen that the larger the inference capability $g_{i,j}^{m,n}$ is, the more downsampling during the mapping from pose ϕ_m to pose ϕ_n is. Hence the inference from ϕ_m to ϕ_n is more reliable and face recognition should perform better. On the other hand, if the $g_{i,j}^{m,n}$ is smaller, it means more upsampling is conducted during the mapping and hence the inference is less reliable. The basic reason underlying this conclusion is that high resolution in the training indicates more confidence in recognition compared to low resolution in training.

By using different pairs of training and test poses, an array of ICM is generated in Figure 10. We can combine it with the Fisher ratio map, such that we have an overall *discrimination power map (DPM)* $p_{i,j}^{m,n}$. This is done using element-by-element multiplication of $g_{i,j}^{m,n}$ and $f_{i,j}^{m,n}$. Based on $p_{i,j}^{m,n}$, we can compute the discrimination power for each training pose similar to (2) and (3). Figure 11 shows both the Fisher ratio and discrimination power results. We can see that by integrating the geometry information, the discrimination power augments the advantage of the 32° pose. This shows that a 32° pose has advantages from geometric point of view, i.e., relatively less upsampling is done when we infer from 32° pose to other poses.

Figure 12 shows the DPM overlaid on the projected triangles for all poses. Basically, for each pose ϕ_m, we assign the intensity of the projected triangle $\mathbf{t}_u^m$ as the discrimination power $p_{i,j}^m$ of its corresponding appearance patch. $p_{i,j}^m$ is simply the average of $p_{i,j}^{m,n}$ along all test poses ϕ_n. On average, the intensity of the 32° and −32° poses is brighter than other poses, as indicated by the curve "DPM: test pose within all poses" in Figure 11. This shows the advantage of the 32° pose in recognition. Furthermore, this plot shows that for each training pose, which patch contributes more to the recognition. Interestingly we see that the facial features, such as the mouth and nose, seem to contribute less to recognition than the cheek. In part this is because that it is relatively hard to register the facial features across multiple poses of the same subject due to geometry variation. Therefore, these parts will have relatively larger intra-subject variation, which lowers the Fisher ratio.

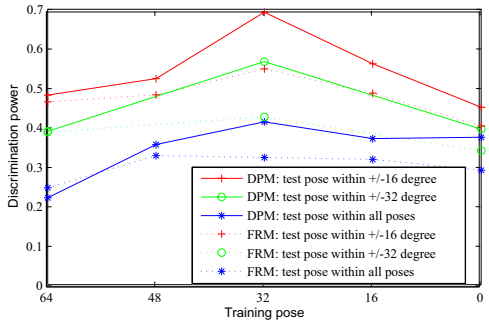

Figure 11. Average discrimination power map (DPM) and average Fisher ratio map (FRM) with respect to different training poses. DPM augments the advantage of 32° over FRM.

5 Conclusions and discussion

If a face recognition system has the freedom to select any pose as the training data, which pose should be chosen in order to achieve the best pose-robust recognition performance? This is both a scientifically interesting and a practically beneficial question. This paper attempts to answer this question and provides justification regarding the answer. First, we conclude that 32° is the optimal training pose by using an experimental approach, where different types of face recognition algorithms are employed and consistent results are observed. Second, we analyze the discrimination power of each pose against all other poses using both appearance and geometry information.

To compare our conclusion with that from psychology, we suggest the two conclusions are somewhat similar. Because both agree that the 3/4 view of learning leads to better recognition than the front and profile views, even though we did not test on exactly 45°, which could be interpreted from nearby 48° and 32°. The two conclusions disagree partly due to the discrete sampling of poses. For example, normally only very coarse poses (such as front, 3/4 or profile view) are studied in psychology. Another reason is the inherent difference between the mechanism of human face recognition and machine face recognition.

Due to limited pose sampling in the database, we consider only discrete poses while looking for the optimal pose. A closer-to-continuous conclusion might be obtained with a database having denser sampling of poses, or synthesized views. One of the future direction might be to rely the analysis more on generic facial geometry information, which is not limited by a particular face database.

We are aware that the optimal training pose of 32° is concluded based on the assumption that the test poses are uniformly distributed. In practical applications, this assump-

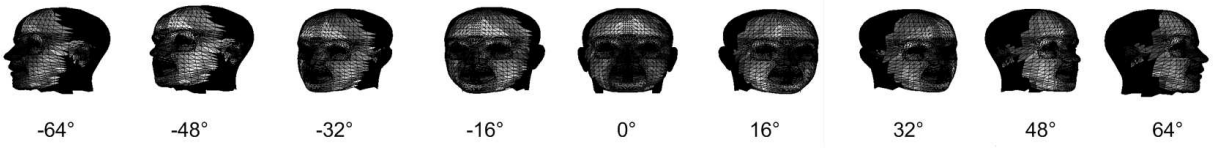

Figure 12. For each training pose, projected triangle mesh overlaid with DPM containing both appearance and geometry information. The brighter regions contribute more to face recognition.

tion might not be true. The actual distribution will depend on the application domain, the position of the camera, etc. However, once the distribution is obtained based on the domain knowledge, the optimal training pose can be found by multiplying each row in Figure 6 with the distribution.

With this paper, we hope to stimulate research into *optimal single-instance recognition* (or opportunistic face recognition [21]). There are many relevant problems along this direction, such as which lighting is the best for face recognition, which gaze is the best for personal identification. These problems are interesting themselves and can help us better understand the intrinsic mechanism of object recognition. On the other hand, these topics have practical values, especially for applications such as Biometrics.

References

[1] P. Belhumeur, J. Hespanha, and D. Kriegman. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(7):711–720, 1997.

[2] V. Blanz and T. Vetter. Face recognition based on fitting a 3D morphable model. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(9):1063–1074, 2003.

[3] V. Bruce, T. Valentine, and A. Baddeley. The basis of the 3/4 view advantage in face recognition. *Applied Cognitive Psychology*, 1:109–120, 1987.

[4] R. Duda, P. Hart, and D. Stork. *Pattern Classification, 2nd edition*. John Wiley & Sons. Inc., New York, 2001.

[5] D. B. Graham and N. M. Allinson. Face recognition from unfamiliar views: Subspace methods and pose dependency. In *Proc. 3rd Int. Conf. on Automatic Face and Gesture Recognition, Nara, Japan*, pages 348–353, 1998.

[6] R. Gross, I. Matthews, and S. Baker. Appearance-based face recognition and light-fields. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 26(4):449–465, 2004.

[7] K. Laugherty, J. Alexander, and A. Lane. Recognition of human faces: Effects of target exposure time, target position, pose position, and type of photograph. *Journal of Applied Psychology*, 55(5):477–483, 1971.

[8] H.-S. Lee and D. Kim. Pose invariant face recognition using linear pose transformation in feature space. In *ECCV 2004 Workshop on Computer Vision in Human-Computer Interaction, Czech Republic*, pages 211–220, 2004.

[9] W.-S. Lee and K.-A. Sohn. Database construction & recognition for multi-view face. In *Proc. 6th Int. Conf. on Automatic Face and Gesture Recognition, Seoul, Korea*, pages 350–355, 2004.

[10] C. Liu and A. Chaudhuri. Reassessing the 3/4 view effect in face recognition. *Cognition*, 83(1):31–48, 2002.

[11] X. Liu and T. Chen. Pose-robust face recognition using geometry assisted probabilistic modeling. In *IEEE Computer Vision and Pattern Recognition, San Diego, California*, volume 1, pages 502–509, 2005.

[12] R. Logie, A. Baddeley, and M. Woodhead. Face recognition, pose and ecological validity. *Applied Cognitive Psychology*, 1:53–69, 1987.

[13] K. Okada and C. von der Malsburg. Pose-invariant face recognition with parametric linear subspaces. In *Proc. 5th Int. Conf. on Automatic Face and Gesture Recognition, Washington D.C., USA*, pages 64–69, 2002.

[14] A. J. O'Toole, T. Vetter, and V. Blanz. Three-dimensional shape and two-dimensional surface reflectance contributions to face recognition: An application of three-dimensional morphing. *Vision Research*, 39:3145–3155, 1999.

[15] S. Palmer, E. Rosch, and P. Chase. *Canonical perspective and the perception of objects*. Number 9 in J. Long and A. Baddeley(Eds), Attention and Performance. Erlbaum Hillsdale, NJ.

[16] P. Phillips, P. Grother, R. Micheals, D. Blackburn, E. Tabassi, and J. Bone. Face recognition vendor test (FRVT) 2002: Evaluation report.

[17] T. Sim, S. Baker, and M. Bsat. The CMU pose, illumination, and expression database. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(12):1615–1618, 2003.

[18] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.

[19] M. Weber, W. Einhuser, M. Welling, and P. Perona. Viewpoint-invariant learning and detection of human heads. In *Proc. 4th Int. Conf. on Automatic Face and Gesture Recognition, Grenoble, France*, pages 20–27, 2000.

[20] D. Weinshall and M. Werman. On view likelihood and stability. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(2):97–108, 1997.

[21] Y. Yacoob and L. Davis. Smiling faces are better for face recognition. In *Proc. 5th Int. Conf. on Automatic Face and Gesture Recognition, Washington D.C., USA*, pages 59–64, 2002.

[22] W.-Y. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Computing Survey*, 35(4):399–458, 2003.