Perspective Symmetry Invariant and Its Applications

Tianqiang Yuan, Shuicheng Yan, Xiaoou Tang
Department of Information Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong

Abstract

Face is a perceptually symmetric object; however, it often appears not so in captured image due to the rotation in depth within the 3D space. In this paper, we explore the invariant of the symmetry point pair when the pose and symmetry radius are changed. Under the general perspective camera model, we present an invariant, called perspective symmetry invariant (PSI), which only depends on the focus length and the rotation angle in depth when the angle is relatively small. PSI is discriminating for pose estimation; meanwhile, it can be used to synthesize face shape in novel views with different PSI values. Encouraging experiments on these two contributions of PSI are presented.

1. Introduction

Pose[1] variation may greatly affect the performance of a face recognition system [6]. Unfortunately, 75% of faces in real captured images are not absolutely frontal [8]; hence the pose estimation becomes an inevitable process in order to implement an automatic face recognition system, and it often acts as face alignment [4] step between the face detection [10] [11] and final face recognition steps.

Previous algorithms for face pose estimation can be roughly divided into two categories: appearance based and geometry based approaches. Appearance based approaches treat pose estimation as a general classification problem [3][9]. For geometry based approaches, the knowledge of projective geometry is applied to explain the relationship between the 3D scene and the 2D image by a set of key feature points [2][6][7]. Geometry based method is relatively more accurate since it is not effected by illumination, expression and age variants; moreover, it does not need the training stage, hence is more applicable. The faces with the same pose in 3D space might not look similar after mapped into 2D image plane, which depends on the position where the faces are mapped in the image, and Fig. 1 displays two such kind of examples. However, most previous geometry based approaches cannot well explain this issue since they often used simplified camera models like weak perspective model [2].

In this paper, we try to resolve the above problem in theory. We study the symmetric point pairs under general perspective camera model, which is more general than weak perspective model. By exploring the relationship between symmetry property and the camera parameters, we deduce a perspective symmetry invariant (PSI), which has the following two properties: 1) for the given rotation angle in depth of a symmetric point pair, PSI is independent to its symmetry radius; and 2) when rotation angle in depth is relatively small, the relationship between 1/PSI and rotation angle in depth can be approximately characterized with a cotangent function. PSI is an effective feature for pose estimation; and the experiments demonstrate the encouraging results for pose estimation with PSI. Moreover, PSI can be used to synthesize face shape of novel views by using different PSI values, which is of potential application importance for constructing non-frontal active shape models from a set of fontal face shapes, which can be used for non-frontal face alignment.

Figure 1. Examples: faces with same pose in 3D space, yet different poses in image plan.

2. Symmetry Analysis for Pose Analysis

2.1. Perspective Symmetry Invariant

In this section, we apply the perspective camera model to explore the properties of the symmetry object rotated in depth. For simplicity, the principle point is assumed to be the center of image plane temporarily. Fig. 2 is the top-view of the camera model from the y axis, and the image plane is simplified as the x-axis. C is the origin of the image coordinate system. The two symmetric points have the same radius, r, to their middle point. These three points are mapped to the image plane as point L, O, and R, respectively. Let f $(f > 0)$ represent the focus length and z $(z > f)$ denote the distance from the middle point to the focus point. When there is no rotation in depth, it is obvi-

[1]The work in this paper was fully supported by grants from the Research Grants Council of the Hong Kong Special Administrative Region.

The 18th International Conference on Pattern Recognition (ICPR'06)
0-7695-2521-0/06 $20.00 © 2006 IEEE

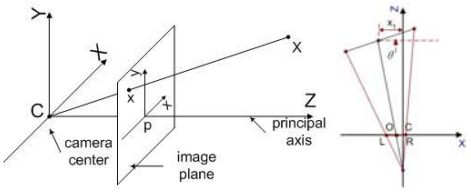


Figure 2. Top-view of symmetric point-pair in perspective camera model.

ous that $\|OL\| = \|OR\|$ ($\|*\|$ is the distance of two points) even there is an offset, x_1 $(-\infty < x_1 < +\infty)$, from the middle point to the Z-axis. Denote θ $(-90^\circ \leq \theta \leq 90^\circ)$ as the rotation angle in depth as shown in Fig. 2. Positive direction means rotating from the line of the points to the horizontal axis clockwise. From the most general situation in Fig. 2 (right), we have

$$\begin{aligned}\|OL\| &= \|CL\| + sign(x_1)\|CO\| \\ &= \frac{f}{z - r\sin\theta}(r\cos\theta - x_1) + \frac{f}{z}x_1 \\ &= \frac{fr(z\cos\theta - x_1\sin\theta)}{z(z - r\sin\theta)},\end{aligned} \tag{1}$$

$$\begin{aligned}\|OR\| &= \|CR\| - sign(x_1)\|CO\| \\ &= \frac{f}{z + r\sin\theta}(r\cos\theta + x_1) - \frac{f}{z}x_1 \\ &= \frac{fr(z\cos\theta - x_1\sin\theta)}{z(z + r\sin\theta)}.\end{aligned} \tag{2}$$

In the above two equations, the numerators of $\|OL\|$ and $\|OR\|$ are the same, so

$$\frac{1}{\|OR\|} - \frac{1}{\|OL\|} = \frac{2z\sin\theta}{f(z\cos\theta - x_1\sin\theta)}. \tag{3}$$

In this equation, the variant r disappears, that is, $\frac{1}{\|OR\|} - \frac{1}{\|OL\|}$ is independent on r. Moreover, z, f, θ, x_1 are fixed when the middle point of the symmetric point-pair is fixed; hence the above equation is constant, and we have the *Perspective Symmetry Invariant*:

$$PSI = \frac{1}{\|OR\|} - \frac{1}{\|OL\|} = \frac{2z\sin\theta}{f(z\cos\theta - x_1\sin\theta)}. \tag{4}$$

Further, if $\theta \neq 0$, we can calculate the reciprocal of PSI as

$$\begin{aligned}\frac{1}{PSI} &= \frac{\|OL\|\|OR\|}{\|OL\| - \|OR\|} = \frac{f}{2}\mathrm{ctg}\theta - \frac{1}{2}\frac{fx_1}{z} \\ &= \frac{f}{2}\mathrm{ctg}\theta - \frac{1}{2}sign(x_1)\|CO\|.\end{aligned} \tag{5}$$

Here, z and x_1 also disappear. $\|CO\|$ is the distance from the mapped middle point to the origin in the image plane. There is only one unknown parameter, f, in the equation, when the the middle of the symmetric pair points is fixed. More generally, when the offset of principal point is considered, we only need to add it to the right side as

$$\begin{aligned}\frac{1}{PSI} &= \frac{f}{2}\mathrm{ctg}\theta - \frac{1}{2}\frac{f(x_1 + x_0)}{z} \\ &= \frac{f}{2}\mathrm{ctg}\theta - (\frac{1}{2}sign(x_1)\|CO\| + C_1),\end{aligned} \tag{6}$$

where x_0 is an offset between the origin and the principle point of camera, $C_1 = \frac{1}{2}\frac{fx_0}{z}$ is a constant. Further, $\mathrm{ctg}\theta$ is very large when θ is small, thus

$$\frac{f}{2}\mathrm{ctg}\theta \gg (\frac{1}{2}sign(x_1)\|CO\| + C_1).$$

By ignoring the item of right side, we can obtain

$$\begin{aligned}PSI &= \frac{1}{\frac{f}{2}\mathrm{ctg}\theta - (\frac{1}{2}sign(x_1)\|CO\| + C_1)} \\ &\approx \frac{1}{\frac{f}{2}\mathrm{ctg}\theta} = \frac{2}{f}\mathrm{tg}\theta.\end{aligned} \tag{7}$$

It shows that, when the rotation angle is small, the perspective symmetry invariant (PSI) can be considered to be independent to the distant $2r$ between two symmetric points, the offset of the center of these two symmetric points, and camera center offset. PSI can be used for many applications, including pose classification and new view synthesis of face shape.

2.2. Face shape synthesis with PSI

From (4), we can synthesize face shapes with different PSI from a frontal template. In Fig. 3 (a), the left side of left eye and mouth are presented, and the right side of the left eye is synthesized from its left side according to the symmetry property, and the rest of the face can be generated with the same method. In this problem, the length of one side l_1 and the PSI are given. The length of the other side l_2 is to be calculated. According to (4),

$$1/l_1 - 1/l_2 = PSI. \tag{8}$$

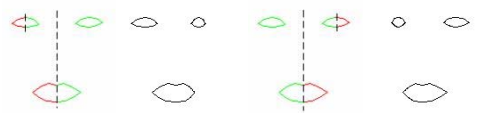
Here, if $PSI > 0$, which means $l_2 > l_1$; when l_1 satisfies that $1/l_1 < PSI$, l_2 will be calculated as a negative value, which is unreasonable. So, we set PSI as a negative value and synthesize the shorter half from the longer half. Thus

$$l_2 = l_1/(1 - l_1 PSI). \tag{9}$$

where l_2 is less than l_1 obviously. Moreover, it is easy to prove that l_2 monotonically increases with l_1 when $PSI >$

The 18th International Conference on Pattern Recognition (ICPR'06)
0-7695-2521-0/06 $20.00 © 2006 IEEE

0; and its limitation is $\lim_{l_1 \to \infty} l_2 = \lim_{l_1 \to \infty} \frac{l_1}{1 - l_1 PSI} = -\frac{1}{PSI}$. These properties are all explainable. Hence, we need to consider the direction of rotation during the pose generation of face shape. As shown in Fig. 3, the face shapes of different poses are synthesized from the red part to the left part, and there is an assumption that face shape is in a plane, and the middle points of the symmetric point-pairs all lie on a line.



(a) (b)

Figure 3. Shape synthesis with different PSI.

3. Experiments

3.1. Relationship Validation between 1/PSI and θ

In this subsection, we explore the relationship between PSI and the rotation angle in depth with the real data, and validate the effectiveness of PSI. Eight frames of different poses are selected from a video as shown in Fig. 4. We label the eye corners manually and their angles are estimated by assuming that the head rotates evenly. Then, we have two symmetric point-pairs for each image and assume their middle points have the same horizontal values.

Denoting the four corner points as x_{l1}, x_{l2}, x_{r2}, and x_{r1} from left to right, we can obtain the following equation to calculate the middle of eyes, m, according to the (9) for each image.

$$\frac{1}{m - x_{l2}} - \frac{1}{x_{r2} - m} = \frac{1}{m - x_{l1}} - \frac{1}{x_{r1} - m}. \quad (10)$$

Let $y = 1/PSI$, we can obtain the following equation with (10) by average the PSI values of two data pairs.

$$2y = \frac{(m - x_{l2})(x_{r2} - m)}{(2m - x_{l2} - x_{r2})} + \frac{(m - x_{l1})(x_{r1} - m)}{(2m - x_{l1} - x_{r1})} \quad (11)$$

Then, we approximate distribution of $1/PSI$ with respect to θ with a cotangent-like curve (with a constant plus), i.e. $y = f/2ctg\theta + b$ (f and b are unknown) in the sense least square error.

Based on the cotangent-like curve, Table 1 illustrates the detailed results of these eight frames; the mean and standard deviation of the absolute error are 6.2380 and 4.1721, respectively. These errors should be mainly caused by the mis-locations of the feature points; when the face rotation angle is too large, the eye corners may be occluded, thus are difficult to be accurately located.

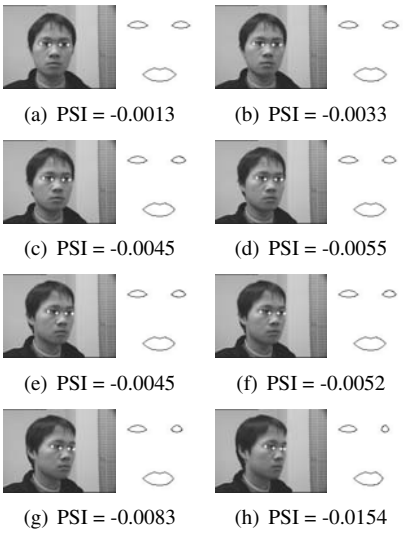
Figure 6 (*)	y	Estimation Angle	Real Angle	Absolute Error
(a)	688.0	7.6	8	0.4
(b)	322.2	20.5	16	3.5
(c)	259.0	28.5	24	3.5
(d)	225.1	35.7	32	3.7
(e)	249.0	30.3	40	9.7
(f)	229.6	34.6	48	13.4
(g)	180.1	51.5	56	4.5
(h)	142.8	73.2	64	9.2

Table 1. Errors of the estimated angles

As described in Section 2.2, we can synthesize face shapes with different PSI values. To synthesize shapes in different size, the PSI needs to be adjusted since

$$1/(\alpha l_1) - 1/(\alpha l_2) = PSI/\alpha. \quad (12)$$

where α is the ratio between the size of real face and the template. Fig. 4 illustrates the simulation results of these 8 frames. The shapes all look like in the real face images. These results provides important heuristic for face alignment with non-frontal pose to construct non-frontal face models from frontal face by rendering with different PSI values.



(a) PSI = -0.0013 (b) PSI = -0.0033

(c) PSI = -0.0045 (d) PSI = -0.0055

(e) PSI = -0.0045 (f) PSI = -0.0052

(g) PSI = -0.0083 (h) PSI = -0.0154

Figure 4. Face shape synthesis with derived PSI from two symmetric point pairs.

3.2. Error Analysis

The limitation of geometry based approaches for pose estimation is that they may be sensitive to the location errors of feature points. In this subsection, we explore the tolerance ability of PSI.

The 18th International Conference on Pattern Recognition (ICPR'06)
0-7695-2521-0/06 $20.00 © 2006 IEEE

Figure 5. Pose estimation results

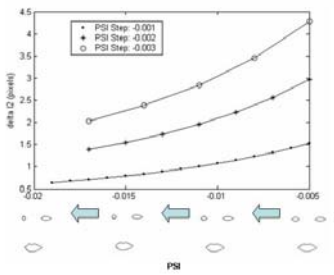


Figure 6. The distribution of ∇l_2 vs. PSI.

Firstly, we assume the term $(\frac{1}{2}sign(x_1)\|CO\| + C_1)$ is omit in (6). In other words, the middle point is on the principle axis. So $PSI = 2\text{tg}\theta/f$ is a tangent function exactly. Because f is still an undetermined parameter, we use PSI to grade the poses instead of rotation angle θ. By setting l_1 to a constant, l_2 can be calculated with (9). Here, we set $l_1 = 50$ pixels. The range of PSI is [-0.005, -0.020] and there are two different step intervals, -0.001, -0.002, and -0.003, for PSI. For each PSI value, from (9), we can obtain two sets of $l_{2,i}, i = 1, 2, ...$ Then, the interval of $l_{2,i}$ can be computed as $\nabla l_{2,i} = l_{2,i+1} - l_{2,i}$.

3.3. Face Pose Estimation with PSI

In this experiment, we collect 297 face images in the UMIST face database [1], which consists of face images of twenty people. Each of them covers a range of poses from profile to frontal views. The four eye-corners are manually labeled and we select 7 images as in Fig. 5 for templates of different poses. From (10), we can obtain their middle points and calculate their PSI as references.

For the rest faces, their PSIs are calculated in the same way. By comparing with the above 7 templates, we classify them to the closest one. Fig. 5 illustrates some of the results. The first face in each cluster is the referent template, and we can see the pose classification results are basically acceptable, even though there may exist mis-locations of the feature points.

The larger is $\nabla l_{2,i}$, the more tolerance ability the related PSI has. Fig 6 plots distribution of $\nabla l_{2,i}$ with respect to different PSI; we can estimate the tolerance ability of error location. Since the relationship between PSI and θ is a tangent-like function, PSI changes much when $|\theta|$ is large; hence we can use the upper curve of Fig. 6 for analysis when $theta$ is large, and use the low one when $theta$ is small. From the result in Fig. 11, we can see that PSI is very sensitive to the pose variant when rotation angle is small.

Further, when $(\frac{1}{2}sign(x_0)\|CO\| + C_1)$ is considered, as mentioned in (7), its effect can be ignored when the rotation angle is small. When the rotation angle is large, it will make PSI change slower; Thus, the tolerance ability of PSI is also reduced. To design a precise face pose estimation system, we need not only a precise detection of feature points, but also a good camera calibration.

References

[1] *http://images.ee.umist.ac.uk/danny/database.html*.

[2] T. Alert. 3-d pose from 3 points using weak perspective. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 16(8):802–808, 1994.

[3] J. Bruske, E. Abraham-Mumm, J. Pauli, and G. Sommer. Head pose estimation from facial images with subspace neural networks. *Proc. of Intl. Neural Network and Brain Conference*, pages 528–531, 1998.

[4] T. Cootes and C. Taylor. Technical report: Statistical models of appearance for computer vision. *http://www.isbe.man.ac.uk/ bim/refs.html*, 2001.

[5] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge University Press, 2003.

[6] Y. Hu, L. Chen, Y. Zhou, and H. Zhang. Estimating face pose by facial asymmetry and geometry. *IEEE Conf. on Automatic Face and Gesture Recognition*, pages 651–656, 2004.

[7] T. Huang, A. Bruckstein, R. Holt, and A. Netravali. Multiple view geometry in computer vision. *Uniqueness of 3D pose under weak perspective: A geometrical proof*, 17:1220–1221, 1995.

[8] A. Kuchinsky, C. Pering, M. Creech, D. Freeze, B. Serra, and J. Gwizdka. Fotofile: A consumer multimedia organization and retrieval system. *Proc. of ACM CHI'99, Conference*, 1999.

[9] S. Li, Q. Fu, L. Gu, B. Scholkopf, Y. Cheng, and H. Zhang. Kernel machine based learning for multi-view face detection and pose estimation. *IEEE Conf. on Computer Vision*, pages 674–679, 2001.

[10] S. Li and Z. Zhang. Floatboost learning and statistical face detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 29(9):1112–1123, September 2004.

[11] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple feature. *IEEE Conf. on Computer Vision and Pattern Recognition*, 1(8-14):511–518, 2001.

The 18th International Conference on Pattern Recognition (ICPR'06)
0-7695-2521-0/06 $20.00 © 2006 IEEE