Fully Automated Facial Symmetry Axis Detection in Frontal Color Images

Xin Chen Patrick J. Flynn Kevin W. Bowyer
Department of Computer Science and Engineering
University of Notre Dame, Notre Dame, IN 46556 USA
{xchen2, flynn, kwb}@nd.edu

Abstract

In this paper, we consider the problem of automatically detecting a facial symmetry axis in what we will call a standard human face image (acquired when the subject is looking directly into the camera, in front of a clean gray background under controlled illumination). Images of this kind are encountered in face recognition scenarios; this detection should facilitate more sophisticated facial image processing. The proposed method is based on GLDH (gray level difference histogram) analysis and consists of three components: (1) the face region detection stage crops an approximate face region out of the background, (2) symmetry detection discovers a vertical axis to optimally bisect the region of interest, assuming it is bilaterally symmetric, and (3) orientation adjustment aligns the angle of the symmetry axis with the orientation of the face. An implementation of the method is described, and results are presented. This detector's robust performance is evidenced by its success finding symmetry axes in more than 7,500 images collected from 600 distinct subjects. One of our method's most noteworthy contributions is that, according to our experimental results, many of the automatically detected axes are more accurate than the reference axes. Our automated detector is a powerful tool because it is not as susceptible to human error as its manual counterpart and, as the first application of its kind, it could potentially serve as a new biometric.

1. Introduction

Facial images are becoming increasingly significant in biometrics research. Only a few researchers have tried to make use of the facial symmetry feature. Quintiliano *et al.* proposed a practical procedure to improve face recognition based on symmetrization and principal component analysis[1]. Symmetrization, in this case, means reconstructing the dark side of the face from the clear side, or calculating the face average with its inverse, thus equalizing the face illumination. Through the symmetrization procedure, the eigenface algorithm, which performs poorly with images collected under uncontrolled illumination conditions, manifests significantly improved performance in testing images acquired under unsuitable illumination conditions, specifically when the face image has one well-illuminated side.

Liu *et al.* defined asymmetry measurements for 3D and 2D facial images based on the mid line passing through the midpoint of the line segment connecting the inner canthus and the philtrum [2][3] [4] [5] [6] [7]. The three points were found in real expression video sequences by hand-selecting three points on the first frame and then automatically tracking the rest of the video frames, with the results verified and corrected afterwards by a human user. They used facial asymmetry in such applications as face identification, facial expression classification and gender discrimination. The symmetry measurement (Y score) quantitatively defined in this paper could be at least as powerful as the asymmetry measurement.

Julesz [8] highlights that the presence of axial symmetry can facilitate spontaneous texture perception by humans. Psychophysical experiments have also been used to study the role of bilateral symmetry in face recognition. Troje *et al.* concluded that the generalization performance of the symmetric view of a face is much better than the generalization of other views. Furthermore, he found that mirror reversal causes few additional errors even if it results in an unrealistic view of the learned face [9] [10]. This research confirmed the ability to identify mirror symmetric patterns used for viewpoint generalization by approximating the view symmetric to the learned view by its mirror-reversed image. Zaki *et al.* examined the recognition of faces at novel orientations[11]. While performance tends to decay as the difference between the study and the test angles increases, an orientation that is symmetric with respect to the study orientation shows strong performance and, in many cases, results prove better than those obtained using a frontal view. Symmetrized faces show surprisingly different patterns of behavior than unsymmetrized faces, despite the fact that many faces are already fairly symmetric.

An automatic facial symmetry axis detector is a valuable tool to explore the use of the bilateral symmetry feature of human faces because it:

1. frees researchers from tedious, manual ground-truth writing in preprocessing the face images;

2. reduces human intervention, making results more objective and accurate;

3. proves more time-cost efficient.

Also, the quantitative symmetry measurement in this research may represent a new biometric which can improve discrimination among different categories of people. Though symmetry is showing promise, little work has been done to detect an image's symmetric axis, a feature whose location is crucial to using symmetry in various applications.

Chetverikov [12] successfully applied the feature-based interaction map (FBIM) on texture (and even non-texture face) pattern symmetry analysis. FBIM is based on spatial gray level difference statistics which describe pairwise pixel interactions. However, FBIM is neither robust nor efficient when using standard face images, on which we can isolate only an approximate face region. Motivated by [12] [13], we propose a new method, which makes use of the statistical characteristics of the gray tone dependence in a more effective way for standard facial images.

2. Experimental Design

We developed a fully automatic facial symmetry detector and used it to determine the axis of symmetry in standard face images. For assessment purposes, we also manually locate the face regions on the same face images by clicking on the centers of each eye and the nose tip, and defining the line passing through the middle point of the interocular line segment (connecting the centers of each eye) and the nose tip as a reference axis of symmetry. Two measurements are used to quantitatively evaluate the performance of the detector. The first, *angle error* θ, is the difference between these two axes, as illustrated by Figure 1. The other metric, *shift* s, is the distance between points A and B in Figure 1. s represents the distance between Point A, the intersection of the reference axis and the interocular line, and point B, the intersection of the detected symmetry axis and the interocular line. The calculations of θ and s are used to assess the detector's performance. We also visually examine the faces bisected by the detected axis to qualitatively evaluate the detector's performance, since in many cases, the facial symmetry detector bisects faces visually better than the reference axis. The data used to obtain the results below was acquired at the University of Notre Dame during the 2003 and 2004 academic years; 7,534 color images from 600 distinct subjects were acquired. Each set consists of two frontal views with different facial expressions. Image acquisitions were held weekly for each subject, and most subjects participated multiple times. The sensor, a Canon Powershot G2 camera, produces an 8-bit color image with a resolution of 2272x1704. Two Smith-Victor A120 lights with Sylvania Photo-ECA bulbs provided studio lighting. The lights

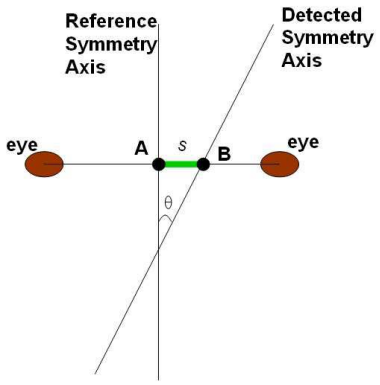
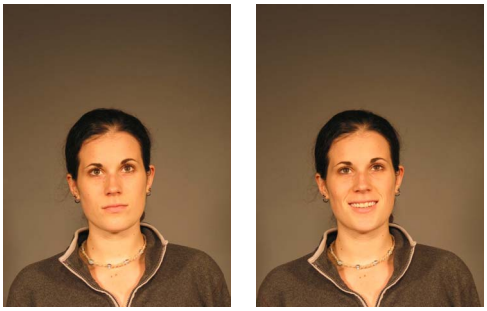


Figure 1: Angle error and shift for evaluation

were located approximately eight feet in front of the subject; they are placed approximately four feet to the subject's left and right. Both lights were trained on the subject's face. The lights were suspended at about six feet; we required all subjects to remove eyeglasses during acquisition. Figure 2 shows one subject's neutral and smiling expressions. Each image features one frontal, vertically-oriented face that may include some of the upper torso and gray background.



(a) Neutral (b) Smiling

Figure 2: Sample face images

3 Face Detection

We first locate the approximate face region since there is significant background, clothes, non-facial skin, and hair in the image. Our facial symmetry detector is robust and does not require that we strictly exclude all non-face pixels. We segment human skin regions from non-skin regions based on a chromatic color model which has been effectively used to segment color images in many applications [14]. Chro-

matic colors are defined as follows:

$$r = \frac{R}{R+G+B}$$

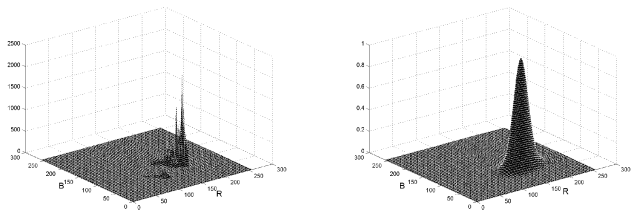
$$b = \frac{B}{R+G+B}$$

A rectangular area on the forehead, chin, and cheek were extracted for each training image, and each one of these rectangular area's pixels was used as one skin sample to determine the color distribution of human skin in the color space. Our samples were taken from thirteen images of thirteen persons of different ethnicities and genders, yielding a variety of facial skin tones. As the skin samples were distilled from the color images, they were filtered using a low-pass filter to reduce the effect of noise in the samples. Figure 3 (a) shows the color distribution of these skin samples in the chromatic color space. The color histogram reveals that the skin-color distribution of different people is clustered in the chromatic color space that can be represented by a Gaussian model $N(m, c)$, where:

$$Mean : m = E(x),\ where\ x = (r\ b)^T$$

$$Covariance : C = E((x-m)(x-m)^T)$$

Figure 3 (b) is the Gaussian skin color model of our training data. With this Gaussian-fitted skin color model, we can



(a) Empirical Skin Color Distribution

(b) Gaussian Skin Color Model

Figure 3: Skin color histograms in chromatic color space

now obtain the likelihood of skin for any pixel in an image. Therefore, if a pixel, having transformed RGB color space to chromatic color space, has a chromatic pair value of (r,b), the likelihood that this pixel represents a subject's skin surface can then be computed [14]. The skin-likelihood image will be a gray-scale image whose gray values represent the likelihood of the pixel belonging to skin. A sample color image and its resulting skin-likelihood image are shown in Figure 4 (a) and (b), respectively. The detected regions may not necessarily correspond to skin. We can only safely conclude that the detected regions have the same color as the subject's skin. Since the skin regions are brighter than the

(a) (b) (c) (d)

Figure 4: Face detection. (a) Original image; (b) Skin likelihood image; (c) Skin segmented image; (d) Final detected face region.

other parts of the images, the skin regions can be segmented from the rest of the image through a dynamic thresholding process [14].The segmented skin graph (binary image) of Figure 4 (a) is shown in Figure 4(c). Not all detected skin regions contain faces. Some correspond to the neck, shoulders and other exposed body parts, while others correspond to non-skin objects (hair, clothes) with colors similar to those of the skin. A skin region is defined as a connected region in the image containing the largest number of skin pixels.

We proceed to determine a more accurate face region using some characteristics from the skin region described above. The procedures are shown below:

1. Compute $d = \frac{\sum_{i=1}^{n} |c_i - c_m|}{n}$, where c_i is the column number of a skin pixel, $c_m = \frac{\sum_{i=1}^{n} c_i}{n}$ and n is number of skin pixels in that region. Given that clothes and hair could make the detected face region wider than it should be, we empirically selected a threshold $2 \times d$ to remove the noise created by hair and clothes, discarding any skin pixel whose $\mid c_i - c_m \mid$ is larger than the threshold.

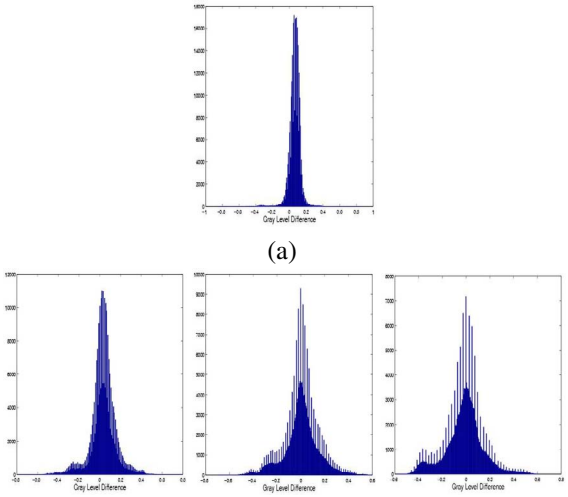
2. We now determine the height and width by moving 4 pointers: one from the left, right, top and bottom of the image. If we find a pixel whose value is not zero, we stop and select it as the coordinate of a boundary. When we have the 4 values, we compute the height by subtracting the bottom and top values and the width by subtracting the right and the left values.

3. The height to width ratio of the face region is approximately one. In the interest of generalization, we set our optimal maximum ratio at 1.5, and opted to eliminate any region below the corresponding height. We determined, by analyzing the results in our experiments, that one would be an appropriate lower limit. We increased the height by extending the lower boundary until the ratio is two, because in most of these cases, the region is very narrow.

Figure 4 (d) shows the final cropped face region, which is ready for symmetry analysis.

4 Symmetry Detection

Given the real symmetry axis on an ideal bilaterally symmetric image, the gray level difference between a face pixel and its symmetric counterpart should be zero. An imaged human face is not perfectly bilaterally symmetric due to intrinsic facial attributes and environmental effects introduced during acquisition. Also, after the face detection stage, we have only an approximate face region, which could include non-facial objects, making it less symmetric. However, gray level difference can still be used as a good indicator of facial symmetry. As illustrated in Figure 5, we first manually locate the most accurate vertical axis which bisects the face evenly. Then we sweep the axis horizontally to the left in twenty-pixel intervals. For each axis location, we display the gray level difference histogram (GLDH). GLDH exhibits greater variance when the axis deviates from the true symmetric axis. In other words, the closer the detected axis is to the true symmetric axis, the denser and more concentrated the GLDH becomes. To describe how wide and loose the GLDH shape is, we employed two measurements, MEAN and gray level difference variance.



(a) (b) (c) (d)

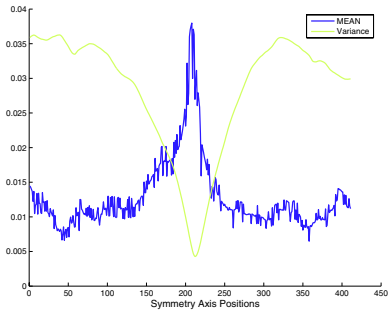
Figure 5: GLDH. (a) Correct symmetry axis; (b) Axis shifted 20 pixels away; (c) Axis shifted 40 pixels away; (d) Axis shifted 60 pixels away.

4.1 MEAN

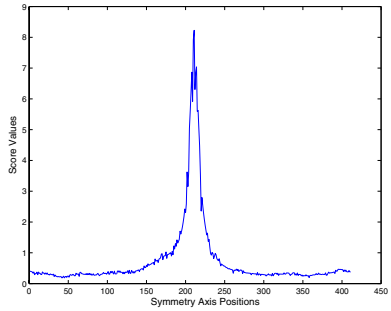
In most standard face regions, given a correct axis estimate, the gray level differences should yield approximately the same value from pixel to pixel. If the axis is incorrectly determined, gray level differences (GLDs) exhibit greater variance. We set a threshold of ± 4 units around the mean. The number of the pixel pairs whose gray level difference falls within the range is defined as MEAN. The higher the MEAN is, the more symmetric the axis. Figure 6 (a) shows the MEANs swiping the axis from left to right on the face region. We observe that, in most cases, the ideal axis lies at, or close to, the peak of MEAN.

4.2 GLD Variance

The variance of gray level difference will become increasingly large relative to the widening GLDH shape. Hence, we show the variance when swiping the axis from left to right in Figure 6 (a). We observe that, in most cases, the ideal axis lies at or close to the minimum of variance. Both



(a) MEAN and variance



(b) Y score

Figure 6: MEAN, variance, and Y score

MEAN and GLD variance are good indicators. In order to combine them, we define the Y score as:

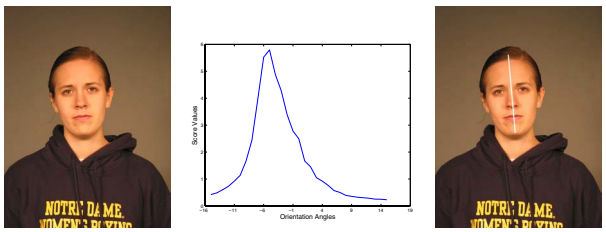
$$Y = \frac{MEAN}{Variance}.$$

The higher the Y value is, the more symmetric the axis is. Figure 6 shows the Y scores sweeping the axis from left to right.

5 Facial Orientation

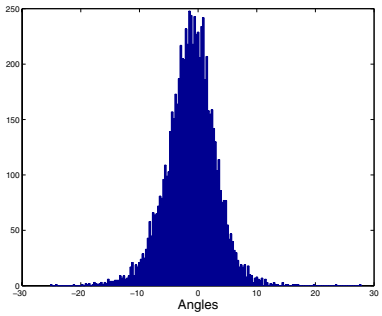
In most cases, the face is not perfectly vertically oriented, either due to the subject's pose during acquisition, or the

position of the camera. Hence, a vertical symmetry axis is not sufficient to evenly bisect the face. The range of the slight rotations exhibited by the face images is represented by $\alpha \in [-15°, 15°]$. For each possible orientation $\alpha \in \{-15°, -14°, -13°, \ldots, 13°, 14°, 15°\}$, we used the method above to detect the symmetry axis and to calculate a Y score for that axis. The rotation that yields largest Y score is selected. Figure 7 (a) and (b) show the original image and the Y score for each possible orientation angle. The peak score appears at $-6°$, which means that the face is rotated six degrees counter-clockwise. The axis which corresponds to that rotation is illustrated by Figure 7 (c).

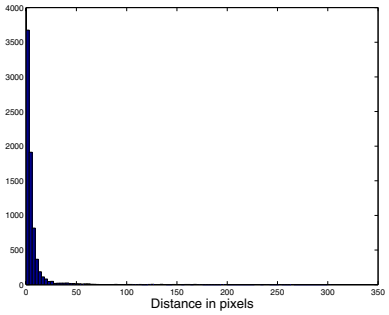


(a) (b) (c)

Figure 7: Orientation adjustment. (a) Original image; (b) Y score for each possible orientation; (c) The axis corresponding to the largest Y value.



(a) Angle error



(b) Shift

Figure 8: Result

6 Results

Figure 8 shows the angle error and shift of the detected symmetry axes for the 7,534 face images. 75.42% of the detected axes are within five degrees difference of the manually detected axes, and 86.86% detected shifts are within 10 pixels distance from the reference symmetry axes. Actually, the detected symmetry axis is more accurate in many images exhibiting large angle error or shift. It is noteworthy that, in a 2272×1704 image, a 10 pixel distance is barely noticeable. In isolated cases, the detector did not yield accurate symmetry axes due to the following factors:

1. Subject exhibits significant head rotation around the y-axis, as shown in Figure 9 (a);

2. Subject wears asymmetric bangs (especially dark hair) that shroud the forehead, as shown in Figure 9 (b);

3. Subject has light hair that is not easily distinguished from skin tone, as shown in Figure 9 (c); Because, in many cases, the detected axes are more accurate than the reference axes, the shift and angle error provide only an approximate performance evaluation. We find that when the symmetry axis goes between the eye centers, it usually yields visually satisfactory results. Seven of 7,534 images featured an incorrectly determined axis despite the fact that there was no significant noise from hair or other interference. We regard these incorrect determinations as error. We may be able to overcome these difficulties if we have more eye location information and restrict our axis search to only the location between the eyes. Figure 10 shows some of the face images whose symmetric axes have been detected.

7 Conclusions

We have developed a fully-automatic facial symmetry axis detector, which can accurately bisect a standard human face image. Fully automatic facial symmetry axis detection is a valuable tool that provides faster and more accurate experimental results than its manual counterpart. Our method defines Y score as the measurement of bilateral symmetry for human faces. This could be a new biometric. We intend to implement this method in analyzing facial images acquired under unstructured lighting conditions, and in acquisition scenarios featuring multiple faces. Because our research's primary focus is automatically detecting the symmetry axis of an isolated face image, face detection is a non-issue, or at most a secondary concern in this paper. Though, there

(a) (b) (c)

Figure 9: Inaccuracy causes

Figure 10: Visual Evaluation

are several mature algorithms that detect the face region in a complex background, more robust face detection algorithms are desired in future applications, especially those that consider facial feature points.

Our method's merit lies in its innovative application of the statistical attributes of GLDH to facial features to simply and efficiently detect the symmetric axis. Our analysis of the GLDH statistical attributes in a facial features context represents a novel approach yielding a straightforward, robust method to automatically detect symmetry. Furthermore, our research is the first to address automatically finding the symmetry axis, and we have, in its course, defined new metrics for facial symmetry and recognition that could potentially expand the frontiers of face recognition research.

References

[1] P. Quintiliano, R. Guadagnin, and A. Santa-Rosa, "Practical procedures to improve face recognition based on eigenfaces and principal component analysis," *The 5th International Conference on Pattern Recognition and Image Analysis: New INformation Technologies*, pp. 372–375, 2001.

[2] Y. Liu, R. Weaver, K. Schmidt, N. Serban, and J. Cohn, "Facial asymmetry: A new biometric," *Tech. Report CMU-RI-TR-01-23, Robotics Institute, Carnegie Mellon University*, August 2001.

[3] Y. Liu, K. Schmidt, J. Cohn, and S. Mitra, "Facial asymmetry quantification for expression invariant human identification," *Computer Vision and Image Understanding*, vol. 91, pp. 138 – 159, July 2003.

[4] S. Mitra and Y. Liu, "Local facial asymmetry for expression classification," *Proc. Computer Vision and Pattern Recognition*, pp. 889–894, 2004.

[5] Y. Liu and J. Palmer, "A quantified study of facial asymmetry in 3d faces," *Proc. IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, p. 222, 2003.

[6] Y. Liu and S. Mitra, "Experiments with quantified facial asymmetry for human identification," *Tech. Report CMU-RI-TR-03-08, Robotics Institute, Carnegie Mellon University*, 2003.

[7] Y. Liu and S. Mitra, "Human identification versus expression classification via bagging on facial asymmetry," *Tech. Report CMU-RI-TR-03-08, Robotics Institute, Carnegie Mellon University*, 2003.

[8] B. Julesz, "Experiments in the visual perception of texture," *Scientific American*, vol. 232, pp. 34–43, 1993.

[9] N. Troje and H. Bulthoff, "How is bilateral symmetry of human faces used for recognition of novel views?," *Vision Research*, no. 38, pp. 79–89, 1998.

[10] N. Troje and H. Bulthoff, "Face recognition under varying poses: the role of texture and shape," *Vision Research*, no. 36, pp. 1761–71, 1996.

[11] T. Busey and S. Zaki, "The contribution of symmetry and motion to the recognition of faces at novel orientations," *Memory and Cognition, in press*.

[12] D. Chetverikov, "Texutre analysis using feature-based pairwise interaction maps," *Pattern Recognition*, vol. 32, pp. 487–502, 1999.

[13] D. Chetverikov, "GLDH based analysis of texture anisotropy and symmetry: an experimental study," *Pattern Recognition*, vol. I, pp. 444–448, 1994.

[14] H. Chang and U. Robles-Mellin *website, http://www-cs-students.stanford.edu/ robles/ee368/main.html*.