Performance Improvement of Automatic Facial Midline Detection by Chain–coded Merlin–Farber Hough Transform

W.Ohyama, D.Okamoto, T.Wakabayashi and F.Kimura

Graduate School of Engineering, Mie University, Tsu-shi, Japan
E-mail: ohyama@hi.info.mie-u.ac.jp

Abstract – **We propose a novel approach for detection of the facial midline from a frontal face image. The facial midline has several applications, for instance reducing computational cost required for facial feature extraction (FFE) and postoperative assessment for cosmetic or dental surgery. The proposed method detects the facial midline of a frontal face from an edge image as the symmetry axis using the Merlin–Faber Hough transformation. Experimental results on the FERET database indicate that the proposed algorithm can accurately detect midlines of face images with different scales and rotation.**

Keyword – **Face Image Processing, Facial Midline, Hough Transform**

I. Introduction

Biometrics employing a fully automatic face recognition or authentication technologies requires both face detection and recognition. In the face detection problem, we are given an input image that may contain one or more human faces (or it may contain no face at all). The scale of the face is not known in advance. The problem is to segment the input image and isolate the face(s). Particularly, it is necessary to determine a tight bounding box around each face that contains just the face (forehead to chin), excluding as much of the hair as possible. Of course, the results of the recognition task[10] depend heavily on how well the detection task has been done.

For a human face, there are important features or landmarks that one can exploit for detection purposes. If the position of these facial features is known, then face detection and localization can be done easily and more accurately.

The detection of facial features, though, is computationally expensive; hence it makes sense to apply the detection only in the vicinity of a face and not the entire image (which may contain many non-face artifacts). Even for frontal face images can be observed as the most simple situation in face recognition, there are many parameters to estimate, for instance location of each feature, scale and rotation of faces. If we get any guides that can be utilized for facial feature extraction by a method that is easier than that for facial features, it is possible to reduce total computational costs.

The facial midline, in other words the facial symmetry axis, has practical applications in the region of facial image analysis. For instance, the midline is one of promising guides to reduce the computational cost in biometrics based on fully automatic face recognition. The extraction of facial midline is equivalent to the detection of facial slant angle and localization of the center point between each eye; hence, the extracted midline can be utilized to normalize the slant and location of the face. This reduces the complexity of facial feature extraction. Also the facial midline has additional contributions for face recognition.

From this background, we proposed a facial midline detector[6, 7] based on the Merlin–Faber Hough transform (MFHT)[5]. This method detects the facial midline from a grayscale image containing one frontal face. Since faces are often slanted in image, the detection method must be robust for these varieties. We define the facial midline as the perpendicular bisector of the interocular line segment (connecting each eyes). Our method is based on bilateral symmetry of human face and extracts the symmetry axis as the facial midline. To extract the axis reliably, we employ MFHT that is able to extract non-analytical curves from an image. For binary images, MFHT behaves an identical algorithm with the correlation method. However it has advantages on computational cost and noise tolerance.

Although the required computational cost of MFHT is smaller than that of the traditional correlation matching method, the cost for voting in the three-dimensional Hough parameter space (x, y, θ) is still not small. To reduce this cost, we restricted the voting to two dimensional parameter space (x, θ) to derive a fast algorithm[7].

In this paper, we present a new performance improvement scheme for midline detection by MFHT. The main concept of the proposed scheme is suppression of redundant vote on the Hough parameter space by introducing chain code representation for the binary edge image.

In contrary to our method, X.Chen *et al.*[2] have proposed an automatic method for the facial midline detection. In their method, axes of facial symmetry are detected as those which maximize the *Y value* that is based on the gray level differences (GLD) between the both sides of the axis. Their approach has the following twofold drawbacks. (1) the *Y value* is quite sensitive to change of lighting conditions: if faces are illuminated from left or right sides, GLD is easily influenced. (2) It is computationally expensive because the maximization problem for the *Y value* is solved by a sweeping algorithm: in other words, to find an axis which maximizes the *Y value*, we have to evaluate all combinations of rotation and position of candidates. Other method has been proposed by Hiremath and Danti[4]. In this method, a face is explained by the *Lines–of–Separability (LS) face model* which includes the facial symmetry axis. To obtain this LS model, we have to extract both eyes from frontal face image before detecting the symmetry axis. From this point of view, this method is observed as a bottom–up approach which is opposite to our method. Another group of methods employ models of facial shape and appearance like *Active Appearance Model*[3]. These methods utilize pre-trained model or template and fit them to input face image. Our proposed method does not require such preliminary training.

This research provides important contribution on the regional innovations. The proposed method does not require any training of the model or classifier. Consequently, the method is able to remove barriers on new entry into the facial image processing by reducing cost required for research and development (R&D).

In the main stream of R&D in facial image processing, we have to involve significant amount of images for improving the performance of algorithms or methods. For instance, a corporation with a market share exceeding 40% employs over 6 million facial images for their R&D of face detection devices. This means that the cost for R&D of face detection devices is significantly increased so that new entrants into this area are strongly blocked. If we have devices embedded with algorithms which do not required such amount of data for training the device, we could encourage competition by new entries.

II. Proposed Methodology

The proposed method consists of three main stages, as shown in Figure 1. In the first stage, we apply preprocessing that consists of edge detection, thresholding and noise removal. The input of the proposed method, demonstrated by (a), is a grayscale image containing one human face in unoccluded frontal view. The size of image is 512×768 pixels. And the face is nonrigid and has a high degree of variability in scale, location, and slant. The resultant image after the preprocessing contains strong edge components of which lengths are sufficient for the MFHT. An example of resultant preprocessed image is shown in (b). The second stage of this method is the MFHT. The MFHT is executed regarding an arbitrary reference point. The reference point is illustrated by $\boldsymbol{p}$ in (b). The MFHT extracts the point that is symmetric to the reference point. The resultant point is called the dual point in this research, which is denoted by $\boldsymbol{q}$ in (c). In the third stage, using the detected coordinates of two symmetric points $\boldsymbol{p}$ and $\boldsymbol{q}$, we obtain the facial midline. Brief descriptions of the each process are presented in [7].

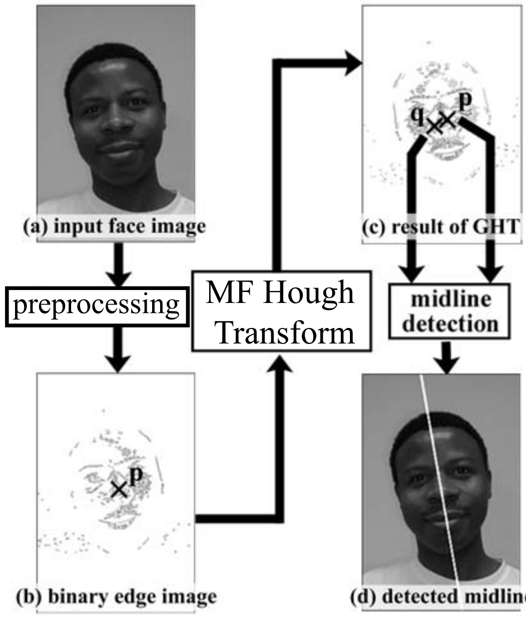


Figure 1. Three main stages of the proposed facial midline detection

A. Merlin–Faber Hough Transform

The Merlin–Faber Hough transform (MFHT) is an algorithm to detect objects, which have the same (or similar) shape as a given template, from given binary images. It is empirically known that MFHT is robust to both noise and lack (or occlusion) of objects in images. For binary images, MFHT behaves as a fast algorithm of template matching. To adapt a template to objects in an input image involving variety of poses: scale, position and rotation, we have to transform and apply the template on the input image by repetition; hence the computation cost becomes high. To reduce this computational time, MFHT employs voting strategy in a parameter space whose dimensionality is equal to the variety of poses.

The MFHT in this research is aimed at finding the dual point that is symmetry to the reference point. The assumption of facial bilateral symmetry suggests that the edge image might also be symmetric. So we employ the mirror image of the binary edge image obtained by preprocess as a template. This means that MFHT detects the most similar shape object to the mirror image from the binary edge image. When MFHT detects the object, we can easily detect the pair of symmetry points.

The tasks of MFHT in the proposed method are as follows:

(1) Selection of the reference point: we select a reference point of MFHT in an image. The selection of the reference point is arbitrary and we have observed that using the center of gravity (CG) of edge pixels as the reference point contributes to the most reliable results[9] .

(2) Generation of the template image: As described

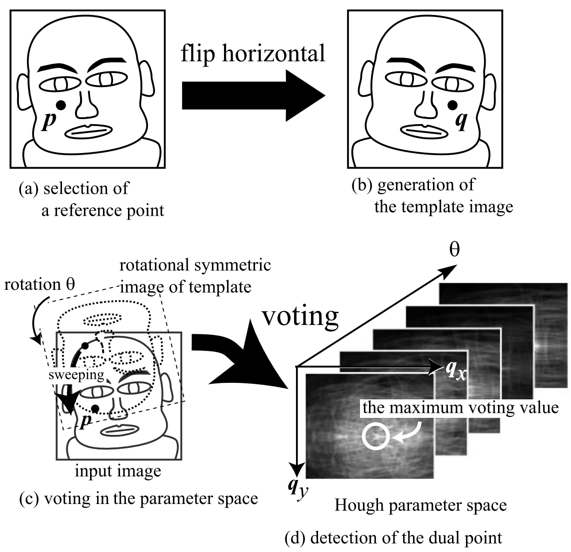


Figure 2. The Merlin–Faber Hough transform in the proposed method

above, we use the mirror image of the binary edge image corresponding to the vertical axis as a template. When the edge image is symmetric corresponding to the vertical axis, the original image and the template overlap completely at the dual point (Figure 2(b)).

(3) Voting in the parameter space: The MFHT's parameter space in this method becomes three dimensional, i.e. q_x, q_y and rotation θ. They correspond to the object's variety of poses. Figure 2(c) illustrates the voting process in this method. The sweeping template, which is point symmetric image of the template (b), scans each of all edge pixels in the binary edge image. During sweeping, the corresponding point in the parameter space accumulates the vote from the template image.

(4) Detection of the dual point: The location and the angle of rotation of the template are detected from the point in the parameter space, where the maximum voting value is obtained (Figure 2(d)).

B. Chain–coded MFHT

Most of the computation time of MFHT is spent for the task of voting in three dimensional Hough parameter space. Reduction of the total amount of the voting contributes substantially to reduce redundant computation time. To reduce the voting amount in MFHT, we propose a new fast algorithm, called Chain–coded MFHT, exploiting the chain code representation of binary edge images.

The detailed tasks of the Chain–coded MFHT is as follows. First, the series of coordinates which represents the binary edge image is converted to be represented by chain code. Similarly, the template image given by the mirror image of the edge image (Figure 2(b)) is represented by chain code independently. Next, in the stage of the sweeping voting (in above subsection (3)), only the edge pixels which have the chain code identical to that of the current edge pixel on sweeping are voted and accumulated on the Hough parameter space. Figure 3 exemplify the difference of resultant Hough parameter space between with and without voting restriction using chain code. From the figure, the amount of voting is reduced substantially by using chain code. If each chain code appears at uniform frequency, the amount of voting on the Hough parameter space is expected to be reduced to one eighth. Computational time for MFHT is reduced by this fast algorithm. In our pilot study, the time for one MFHT operation is reduced from 0.32[s] to 0.23[s] on 2.6 GHz Intel Core2 processor.

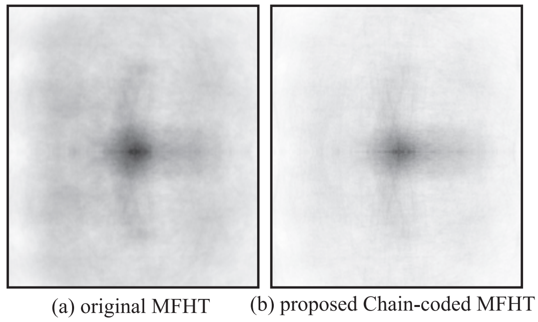


Figure 3. Restriction of voting by Chain–coded MFHT

III. Experiments

To verify the effectiveness of the proposed method, we apply the proposed method to the images from FERET database. Three examples of detected midline are shown in Figure 4. The white line in each picture is the detected midline. The midline of faces over many different scales and rotation has detected correctly.

Next, we quantify the performance of the proposed method by evaluation experiment with 2409 frontal face images from the fa and fb probes in FERET database. For this test, we compare the detected midline with the reference midline obtained from ground-truth eye locations. As used in [2], two measurements, *angle error* $\Delta\theta$ and *distance error* s, are used to evaluate the performance of midline detection. The angle error $\Delta\theta$ is the difference between detected and reference midlines. The distance error s is the distance between theses two midlines on the interocular line segment. Figure 5 illustrates these measures.

To demonstrate the advantage of our proposed method, we compare the proposed method and our former method[7], for the angle and the distance errors. Note that the performance and significance of our former method on midline detection compared to the Chen's GLD–based method[2] have been verified on the same experiments in [6, 7]. Figure 6

Figure 4. Result of the midline detection

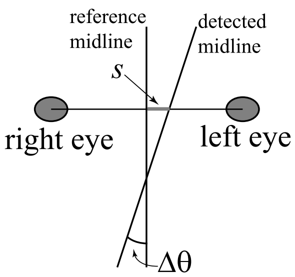
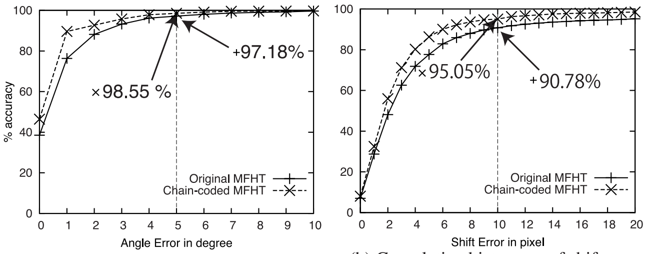


Figure 5. Angle error and distance error for evaluation



(a) Cumulative histogram of angle error (b) Cumulative histogram of shift error

Figure 6. Performance evaluation by cumulative histograms

shows cumulative histograms of the angle error and distance (shift) error of the detected midlines for the 2409 images in FERET by our former method (original MFHT) and the proposed method (Chain–coded MFHT).

98.55% of the detected midlines are within 5 degrees angle error; this means that the rotations of face in 98.55% of input images are correctly estimated by the proposed method. And 95.05% detected distance error are within 10 pixels; this means that the positions of midline in 95.05% of input images are detected correctly. This result suggests that the proposed method provides acceptable performance for the midline extraction. The computational time of the proposed method for one facial images is 0.23[s] by a 2.66 GHz Intel Core2 CPU.

IV. Conclusions

In this paper, we propose a detection methodology for the face midline from an image. Our method based on the MFHT is fast, easy to implement and has good performance. Using detected midlines as a guide for facial feature extraction reduces the computational cost.

Our future work consists of (1) further improvement of the performance, (2) comparing the performance of this method with other methodologies and (3) development of proper application of the detected midline.

References

[1] L. Chen, H. Liao, J. Lin, and C. Han. Why recognition in a statistics-based face recognition system should be based on the pure face portion: a probabilistic decision-based proof. *Pattern Recognition*, 34(7):1393–1403, 2001.

[2] X. Chen, P. Rynn, and K. Bowyer. Fully automated facial symmetry axis detection in frontal color images. In *Fourth IEEE Workshop on Automatic Identification Advanced Technologies, 2005*, pages 106–111, 2005.

[3] G. Dedeoglu, T. Kanade, and S. Baker. The asymmetry of image registration and its application to face tracking. *IEEE transactions on pattern analysis and machine intelligence*, 29(5):807–823, 2007.

[4] P. Hiremath and A. Danti. Detection of multiple faces in an image using skin color information and Lines-of-Separability face model. *International Journal of Pattern Recognition and Artificial Intelligence*, 20(1):39–62, 2006.

[5] P. Merlin and D. Farber. A parallel mechanism for detecting curves in pictures. *IEEE Transactions on Computers*, 100(24):96–98, 1975.

[6] N. Nakao, W. Ohyama, T. Wakabayashi, and F. Kimura. Automatic detection of facial midline as a guide for facial feature extraction. In *Proceedings of the 7th International Workshop on Pattern Recognition in Information Systems*, pages 119–128, 2007.

[7] N. Nakao, W. Ohyama, T. Wakabayashi, and F. Kimura. Automatic detection of facial midline and its contributions to facial feature extraction. *Electronic Letters on Computer Vision and Image Analysis*, 6(3):55–66, 2008.

[8] P. Quintiliano, R. Guadagnin, and A. Santa-Rosa. Practical procedures to improve face recognition based on eigenfaces and principal component analysis. In *The 5th International Conference on Pattern Recognition and Image Analysis: New information Technologies*, volume 11, pages 372–375, 2001.

[9] M. Sato, H. Ogawa, and M. Ookayama. OPTIMIZATION OF THE GENERALIZED HOUGH TRANSFORM. In *Proceedings on 4-th International Conference on Image analysis and processing*, page 221. Plenum Publishing Corporation, 1988.

[10] W. Zhao, R. Chellappa, P. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *Acm Computing Surveys (CSUR)*, 35(4):399–458, 2003.