ACCV 2002: The 5th Asian Conference on Computer Vision, 23 – 25 January 2002, Melbourne, Australia.

Morlet Wavelet Based Human Head and Face Boundary Extraction

Liming Zhang *and* Patrick Lenders
School of Mathematical and Computer Sciences
University of New England
Armidale, NSW 2351, Australia
Email: {liming, pat}@mcs.une.edu.au

Abstract

This paper presents a novel approach to extract human head boundary in plain background image using $2D$ Morlet continuous wavelet transform. Continuous wavelet transform and discrete wavelet transform have their specific properties so that their specific field of applications. In this paper we utilize the rich edge information provided by the continuous wavelet transform in various directions to extract the head boundary, and then fuse those information to get the final result. An algorithm for face boundary extraction is also described in this paper. Effectiveness of the proposed method is shown by our tests with Yale Face Database.

1. Introduction

Automatic recognition of human faces plays an important role in many applications, such as personal identification, security systems and human computer interaction [1]. Face detection and segmentation are the first and most important steps for face image processing and face recognition. Location and description of head boundary are the key problem in face segmentation. The head boundary also has a significant effect in model-based image coding [2] and some of the facial feature extraction approaches [3]. The face boundary has been used in face recognition as part of an extended feature vector [4].

Wavelet transform has come to be an important tool of mathematical analysis, with a wide and ever increasing range of applications, in recent year. A large amount of literature [7, 8] have shown that wavelet transform has certain important inherent generic advantages and so are nearly optimal for a wide class of problems. The wavelet transform used in computer vision is a relatively new concept – about 10 years old, but yet there are quite a number of articles written using this technique. Most of the papers are in the area of image compression and reconstruction. Few paper can be found in the area of human face recognition. In [9] and [10], Gabor continuous wavelet transform has been used for facial feature extraction. In [5] discrete wavelet transform – Haar wavelet transform has been used for facial edge detection.

In this paper, we propose a novel approach to extract the head and face boundaries using Morlet wavelet, a continuous wavelet transform. The 2D continuous wavelet transform based on Euclidean group (consisting of rotation and translation) with dilation has the potential to analyze the directional features of the images [6]. We are not aware of any previous work that has used the directional effect of the continuous wavelet transform in the area of face recognition for edge extraction. In this paper, we use the 2D Morlet wavelet to extract the head boundaries in four directions. The edge images can complement with each other and nearly unbroken boundaries can be composed by using these images.

This paper is organized as follows. In Section 2, the $2D$ Morlet wavelet transform based edge extraction is described. The algorithms of the head and face boundary extraction are introduced in Section 3. Section 4 presents the experimental results and discussion. Conclusion is drawn is Section 5.

2. Wavelet Transform in Edge Detection

2.1 Continuous Wavelet Transform

A wavelet is a complex-valued function $\Psi \in L^2(\mathbf{R}^2)$ (abbreviated as L^2 in the rest of the paper) satisfying the so-called *admissibility condition*:

$$\int \frac{|\hat{\Psi}(\omega)|^2}{|\omega|^{-2}} d\omega < \infty, \qquad (1)$$

where $\hat{\Psi}$ is the Fourier transform of Ψ and $\mid \omega \mid = \sqrt{\omega_1^2 + \omega_2^2}$. If Ψ is smooth enough at the origin, then the admissibility condition implies, as a necessary condition,

$$\hat{\Psi}(0) = 0. \tag{2}$$

If Ψ further belongs to L^1, then the definition of Fourier transform implies that (2) is equivalent to

$$\int \Psi(\mathbf{x}) d\mathbf{x} = 0, \tag{3}$$

where $d\mathbf{x} = dxdy, \mathbf{x} = (x, y)$. Under mild conditions on the function Ψ one can show that the conditions (2) or (3) is also sufficient for obtaining (1).

Let f be a two-dimensional signal (image) of finite energy, i.e. $f(\mathbf{x}) \in L^2$, and sometimes $f \in L^2 \cap L^1$. The general continuous wavelet transform of f is defined by

$$W_f(a, \mathbf{b}) = \int f(\mathbf{x}) \overline{\Psi_{a,\mathbf{b}}(\mathbf{x})} d\mathbf{x}, \tag{4}$$

where $0 < a \in \mathbf{R}$, and $\mathbf{b} \in \mathbf{R}^2$.

The transform only uses the scaling and translating effects on the wavelet function Ψ.

Consider $\Psi_{a,\theta,\mathbf{b}}$ as a function on the Euclidean group with dilation, where $\theta \in [0, 2\pi)$. Then

$$\Psi_{a,\theta,\mathbf{b}}(\mathbf{x}) \equiv a^{-1} \Psi(a^{-1} \rho_\theta(\mathbf{x} - \mathbf{b})), \tag{5}$$

where ρ_θ is the rotation around the origin by angle θ in the anti-clockwise direction, with representation

$$\rho_\theta(\mathbf{x}) = (x \cos\theta - y \sin\theta, x \sin\theta + y \cos\theta).$$

The wavelet transform of $f(\mathbf{x})$ based on the Euclidean group is,

$$\begin{aligned} W_f(a, \theta, \mathbf{b}) &= \int f(\mathbf{x}) \overline{\Psi_{a,\theta,\mathbf{b}}(\mathbf{x})} d\mathbf{x} \\ &= a^{-1} \int f(\mathbf{x}) \overline{\Psi(a^{-1} \rho_\theta(\mathbf{x} - \mathbf{b}))} d\mathbf{x}. \end{aligned} \tag{6}$$

2.2 Edge Detection Using $2D$ Morlet Wavelet Transform

The Morlet wavelet is given by the following [6]:

$$\Psi(\mathbf{x}) = e^{i\mathbf{k_0}\cdot\mathbf{x}} e^{-\frac{1}{2}\mathbf{x}\cdot A\mathbf{x}} - e^{-\frac{1}{2}\mathbf{k_0}\cdot B\mathbf{k_0}} e^{-\frac{1}{2}\mathbf{x}\cdot A\mathbf{x}}, \tag{7}$$

where A is a 2×2 positive definite matrix, $B = A^{-1}$, $\mathbf{k_0}, \mathbf{x} \in \mathbf{R}^2$, and $\mathbf{k_0} \cdot \mathbf{x}$ denotes the ordinary inner product of $\mathbf{k_0}$ and $\mathbf{x}$, and so on. $\mathbf{k_0}$ is a fixed parameter indicating the direction of propagation of the wavelet.

In the space of Fourier transform:

$$\begin{aligned} \hat{\Psi}(\mathbf{k}) = & \quad (detB)^{\frac{1}{2}} e^{-\frac{1}{2}[(\mathbf{k}-\mathbf{k_0})\cdot B(\mathbf{k}-\mathbf{k_0})]} \\ - & \quad e^{-\frac{1}{2}\mathbf{k_0}\cdot B\mathbf{k_0}} e^{-\frac{1}{2}\mathbf{k}\cdot B\mathbf{k}}. \end{aligned} \tag{8}$$

The second term in (7), and that of (8) accordingly, is the correction term that guarantees that Ψ satisfies the admissibility condition (1). If $\mid \mathbf{k_0} \mid$ is chosen large, i.e. $|\mathbf{k_0}| \geq \pi(\frac{2}{\ln 2})^{\frac{1}{2}} \cong 5.34$ [6], then the first term of (8) is close to zero for $\mathbf{k} = 0$ that means that (2) is nearly satisfied without the second term. In practice the second term (i.e. the correction term) is negligible. In the application that follows, we will take $A = I$ and drop the correction term. Thus our (truncated) Morlet wavelet reads[10]:

$$\Psi_M(\mathbf{x}) = e^{i\mathbf{k_0}\cdot\mathbf{x}} e^{-\frac{|\mathbf{x}|^2}{2}}, \tag{9}$$

$$\hat{\Psi}_M(\mathbf{k}) = e^{-\frac{|\mathbf{k}-\mathbf{k_0}|^2}{2}}. \tag{10}$$

Based on (6) and (9), we can get

$$\begin{aligned} &W_f(a, \theta, \mathbf{b}) = \\ &\quad a^{-1} \int\int f(x, y) \overline{\Psi_M(a^{-1}\rho_\theta(\mathbf{x} - \mathbf{b}))} dxdy, \end{aligned} \tag{11}$$

where $\mathbf{b} = (b_1, b_2)$ and

$$\Psi_M(a^{-1}\rho_\theta(\mathbf{x} - \mathbf{b})) = e^{i\frac{\mathbf{k_0}}{a}\cdot\rho_\theta(\mathbf{x}-\mathbf{b})} e^{-\frac{1}{2a^2}|\mathbf{x}-\mathbf{b}|^2}.$$

Based on the fact that rotation preserves inner product:

$$\mathbf{k_0} \cdot \rho_\theta(\mathbf{x} - \mathbf{b}) = \rho_{-\theta}\mathbf{k_0} \cdot \rho_{-\theta}\rho_\theta(\mathbf{x} - \mathbf{b}) = \rho_{-\theta}\mathbf{k_0} \cdot (\mathbf{x} - \mathbf{b}),$$

the rotation of $\mathbf{x}-\mathbf{b}$ by an angle θ in the **anticlockwise** is of the same effect as the rotation of $\mathbf{k_0}$ by the angle θ in the **clockwise**. So we can pre-set the direction $\mathbf{k_0}$. Our method consists of the following steps.

- **(i)** Choose a small parameter a. Our experiments in below show that with $a = 4$ the edge extraction has good enough performance.
- **(ii)** Choose a vector $\mathbf{k_0}$ of a large modulus and fix it all the time. Based on the fact that $\mathbf{k}$ can rotate to $(0, k_0)$, in below we choose $\mathbf{k_0} = (0, k_0)$ with $k_0 = 6$.
- **(iii)** Choose $\theta_i = 0°, 45°, 90°$ and $135°$, $i = 1, 2, 3, 4$, respectively, which are suitable for detecting head boundary contours being orthogonal (or nearly) with the directions $90°, 135°, 0°$ and $45°$, respectively. Combining those edge images in four directions, nearly unbroken head edge can be detected.

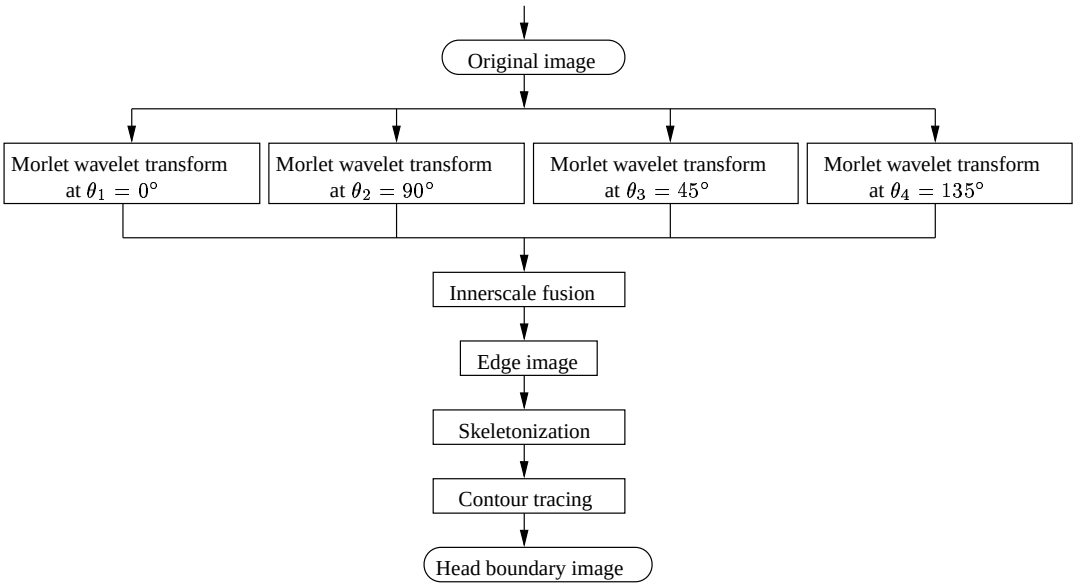

Figure 1: Diagram for head boundary extraction

Having chosen the parameters, the rotated wavelet function is read as

$$\begin{aligned} & \Psi_M(\frac{1}{4}\rho_{\theta_i}(x-b_1, y-b_2)) \\ = \quad & e^{\frac{3}{2}i[(x-b_1)\sin\theta_i + (y-b_2)\cos\theta_i]} e^{-\frac{1}{32}[(x-b_1)^2+(y-b_2)^2]}, \end{aligned}$$

where $i = 1, 2, 3, 4$. Then the Morlet wavelet transform becomes:

$$W_f(a, \theta_i, \mathbf{b}) = \frac{1}{4}\int\int f(x,y)\Psi(\frac{1}{4}\rho_{\theta_i}(x-b_1, y-b_2))dxdy. \quad (12)$$

The discrete formate of (12) is

$$W_f(a, \theta_i, \mathbf{b}) \approx \frac{1}{4}\sum_{x=b_1-\frac{N}{2}}^{b_1+\frac{N}{2}}\sum_{y=b_2-\frac{N}{2}}^{b_2+\frac{N}{2}} f(x,y) \Psi_M(\frac{1}{4}\rho_{\theta_i}(x-b_1, y-b_2)). \quad (13)$$

In our application, the choice $N = 16$ corresponds to a modest computation cost.

3. Head and Face Boundary Extraction

3.1 Head Boundary Extraction

Head boundary refers to the outermost boundary of the head, while the face boundary refers to the face contour excluding the hair, shoulders and the neck. The flowchart of the algorithm for head boundary detection is shown in Fig.1.

(a) (b)

Figure 2: Morlet wavelet transform based edge extraction. (a) Original image. (b) Edge image.

Theoretically, signals are represented with a number of large wavelet coefficients. Noise is evenly distributed across wavelet coefficients and is generally small [7]. Thresholding technique is used to select the wavelet coefficients. First the modulus of the transform coefficients are computed and then the maximum of the coefficients T is found. Denote by B_i the binary image after wavelet coefficient thresholding with different θ_i. The B_i satisfy the following relation:

$$B_i(x,y) = \begin{cases} 0 & |W^{\theta_i}_{\Psi_M}(a, \theta_i, \mathbf{b})| \geq T - C \\ 255 & |W^{\theta_i}_{\Psi_M}(a, \theta_i, \mathbf{b})| < T - C, \end{cases} \quad (14)$$

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

Figure 3: Binary images of Fig.2(a) in four directions after applying Morlet wavelet. (a) $\theta_1 = 0°$. (b) $\theta_2 = 90°$. (c) $\theta_3 = 45°$. (d) $\theta_4 = 135°$.

(a) (b) (c)

Figure 4: Edge extraction through the skeletonizing and contour tracing algorithm. (a) Edge image after skeletonizing algorithm. (b) Head boundary after contour tracing algorithm. (c) Head boundary image.

where $i = 1, 2, 3, 4$. $C = 30$ is based on our experiments. Fig.3 illustrates the binary images in four directions after applying the Morlet wavelet. The original image is shown in Fig.2(a). Fig.2(b) is the edge image after fusing the four images shown in Fig.3. From the edge image we can see the method we proposed can retain unbroken head edges, though there are some noises in the image. The Hilditch skeletonizing approach [11] is used to obtain the single pixel wide boundaries, as shown in Fig.4(a). Contour tracing algorithm is then used to get the location of the boundary points in the image coordinate system and at the same time delete the noises. Fig.4(b) shows the image after performing contour tracing algorithm. The head boundary image of Fig.2(a) is shown in Fig.4(c).

3.2 Face Boundary Extraction

The face boundary is more difficult to be detected due to the poor edge contrast in the face boundary. The flowchart of the algorithm for face boundary extraction is shown in Fig.5.

(a) Inner hair boundary extraction

In this step the hair region is detected using histogram thresholding technique. In the histogram, the first peak area represents the dark regions of the image, such as hair, eyes, eyebrows and mouth. Choose an intensity value between the first two peak areas as the threshold. The binary image $B(x, y)$ consists of most pixels in the hair region. Fig.6(a) demonstrates the result after thresholding. Contour tracing algorithm is then used to get the outlines of each dark regions whose areas are larger than a preset value ρ. Fig.6(b) shows the image after performing contour tracing algorithm. The largest region in $B(x, y)$ is hair region. Then the inner hair boundary can be extracted. Fig.6(c) shows the image of inner hair boundary.

(b) Cheek boundary location

According to the hairstyles, two face models can be constructed. One corresponds to the short hairstyle that the bottoms of the hair are above the ears as shown in Fig.8(b). Another corresponds to the long hairstyle that the bottoms of the hair are below the ears as shown in Fig.2(a). To the long hairstyle, the cheek boundary is included in the inner hair boundary. To the short hairstyle, the following approach is used to locate the

(a) (b) (c)

Figure 6: Inner hair boundar extraction. (a) Binary image. (b) Hair boundary after contour tracing algorithm. (c) Inner hair boundary.

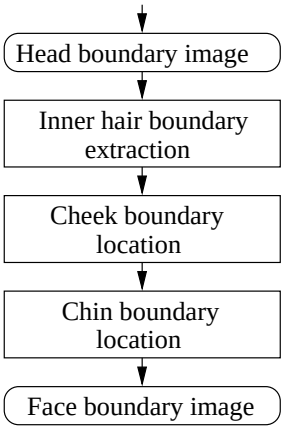

Figure 5: Diagram for face boundary extraction

cheek boundary.

Assume that (h_{l_x}, h_{l_y}), (h_{r_x}, h_{r_y}) represent the inner left and right hair bottom points, respectively, which can be located from the inner hair boundary. From the head boundary and hair boundary, it is easy to detect ears' position where they are outside the inner hair bottom points. Denote the left and right ear bottom points as (x_l, y_l), (x_r, y_r), respectively. Straight line segments between (h_{l_x}, h_{l_y}) and (x_l, y_l), (h_{r_x}, h_{r_y}) and (x_r, y_r) are used to approximate the left and right boundaries for the cheek.

(c) Chin boundary location

In the face image, the contrast at the chin is usually very low (very weak edges) [5]. The edge detecting result in this part can not be relied. In our approach, we use a segment of ellipse to approximate the chin shape.

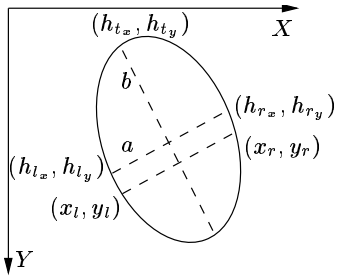

Figure 7: Ellipse model

The ellipse model that is used to approximate the chin shape is shown in Fig.7. The equation for it is:

$$[(x-x')+(y-y')k]^2 + \frac{[(y-y')-(x-x')k]^2}{\rho^2} = \frac{l^2(1+k^2)}{4} \quad (15)$$

where,

$$\begin{cases} k = (h_{r_y} - h_{l_y})/(h_{r_x} - h_{l_x}) \\ x' = (h_{l_x} + h_{r_x})/2 \\ y' = (h_{l_y} + h_{r_y})/2 \\ l = \sqrt{(h_{r_x} - h_{l_x})^2 + (h_{r_y} - h_{l_y})^2} \\ \rho = b/a \end{cases}$$

a and b are the radiuses on Y and X axis, respectively. (h_{t_x}, h_{t_y}) is the top point of the inner hair boundary. The face boundary detected for Fig.2(a) is illustrated in Fig.8(a).

4. Experiment Results and Discussion

The Yale Face Database was used in the experiments. The database contains front view images of 15 people.

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

Figure 8: Examples of results

For each person there are 11 grey level images with different expressions. Out of these 11 images, the head boundary extraction algorithm was successful for 9 images. The algorithm was not effective for 2 images due to one sided light conditiond in the background. Face boundary extraction algorithm relies on the inner hair boundary location to some extent. It is not applicable to the baldheads. There were 117 images processed by the face boundary detection algorithm and about 84% boundaries were picked out. Fig.8 shows some of the result images.

5. Conclusion

In this paper, we propose a novel head boundary extraction approach which involves use of $2D$ continuous wavelet transform. Directional information of the $2D$ continuous wavelet enables the system to extract singularities of images in various directions. The edge detected by using wavelet transform in different directions can compliment with each other and can get back the lost edge information. It can solve the problem where the standard boundary detection becomes ineffective due to edge breaking. Our future task is to find additional information near the chin and accordingly adjust the ellipse segment to get more accurate face boundary.

References

[1] R.Chellappa, C.L.Wilson and S.Sirohey, "Human and Machine Recognition of Faces: A Survey". *Proceedings of IEEE*. Vol. 83. pp.705-740. 1995.

[2] J.B.Waite and W.J.Welsh, "Head boundary location using snakes", *Br. Telecom Technology Journal*, vol.8, pp.127-136, 1990.

[3] C.L.Huang and C.W.Chen, "Human facial feature extraction for face interpretation and recognition", *Pattern Recognition*, vol.25, pp.1435-1444, 1992.

[4] X.Jia and M.S.Nixon, "Extending the feature vector for automatic face recognition", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.17, pp.1167-1176, 1995.

[5] T.Kondo and H.Yan. "Automatic Human Face Detection and Recognition Under Non-uniform Illumination". *Pattern Recognition.* Vol.32. pp.1707-1718. 1999.

[6] J.P.Antoine, R.Murenzi and B.Piette, "Image Analysis with $2D$ Continuous Wavelet Transform: Detection of Position, Orientation and Visual Contrast of Simple Objects". *Proceedings of the International Conference.* France. pp.144-159. 1989.

[7] S.Mallat, "A wavelet tour of signal processing", *Academic Press*, San Diego, 1998.

[8] C.S.Burrus, R.A.Gopinath and H.Guo, "Introduction to Wavelets and Wavelet Transforms: A Primer". Prentice Hall. 1998.

[9] B.S.Manjunath, R.Chellappa and C.Malsburg, "A Feature Based Approach to Face Recognition". *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition.* pp.373-378. 1992.

[10] S.T.McKernna, S.Gong, R.P.Wurtz, J.Tanner and D.Banin, " Tracking Facial Feature Points with Gabor Wavelets and Shape Models". *Proceedings of 1st International Conference on Audio- and Video-Based Biometric Person Authentication, Lecture notes in Computer Science.* Springer Verlag. 1997.

[11] E.R.Davies, "Machine Vision", *Academic Press*. 1997.