Research Article

Arabidopsis Growth Simulation Using Image Processing Technology

Junmei Zhang, Qiuhong Ke, and Ye Tian

School of Technology, Beijing Forestry University, Beijing 100083, China

Correspondence should be addressed to Junmei Zhang; joyzhangjm@163.com

This paper aims to provide a method to represent the virtual *Arabidopsis* plant at each growth stage. It includes simulating the shape and providing growth parameters. The shape is described with elliptic Fourier descriptors. First, the plant is segmented from the background with the chromatic coordinates. With the segmentation result, the outer boundary series are obtained by using boundary tracking algorithm. The elliptic Fourier analysis is then carried out to extract the coefficients of the contour. The coefficients require less storage than the original contour points and can be used to simulate the shape of the plant. The growth parameters include total area and the number of leaves of the plant. The total area is obtained with the number of the plant pixels and the image calibration result. The number of leaves is derived by detecting the apex of each leaf. It is achieved by using wavelet transform to identify the local maximum of the distance signal between the contour points and the region centroid. Experiment result shows that this method can record the growth stage of *Arabidopsis* plant with fewer data and provide a visual platform for plant growth research.

1. Introduction

As an important model plant, *Arabidopsis* has been widely used in plant genetic research [1]. The genome sequencing work was completed at the end of 2000 [2]. To analyse the function of genes, the phenotypes also need to be studied. The phenotypes are the observable characteristics of the plants. They are the inevitable results of the genetic variation and the growth environment [3]. The analysis of plant phenotypes can reveal the impacts of the genes and the environment, which is an efficient method to cultivate new crops.

The phenotypes of *Arabidopsis* include the total area, the number of leaves, and the overall contour. The total area and the number of leaves of a plant reflect its growth situation. The changing of these two parameters can also reveal the growth rate. The overall contour is a significant shape parameter. It can affect the morphology, growth orientation, density, and even size of a plant. With the analysis of the overall contour, we can get a better grasp of the morphological features of the plant.

However, the number of contour points is quite large because of the high resolution of image. It will require much

storage to directly store the contour pixels. Besides, the plant varies at different growth stages, resulting in different numbers of contour pixels. And it will also add difficulties to store the data for their different lengths.

So how to represent the shapes of different *Arabidopsis* plants? Previous research has shown various ways to describe shape variation [4]. What we have done was using elliptic Fourier descriptors (EFDs) [5] to analyse the data in the frequency domain, instead of the spatial domain. It decomposes a curve into a set of harmonically related ellipses [6–9]. EFDs can be easily normalized to represent shapes unaffected by their scale, location, or orientation. It has been widely used for morphological estimation. Menesatti et al. [10] used the Fourier ellipse analysis to extract features of hazelnuts to estimate the shape. Costa et al. [11] extracted the polar signature of the contour of oranges and calculated the coefficients of Fourier harmonics to describe the size, shape, and orientation of each fruit. Similar work has been carried out by other researchers [12–14].

In this paper, EFDs will be used to analyse the contour, and each *Arabidopsis* plant will be represented with much fewer coefficients, rather than the original contour data. This

method not only saves much storage, but also can be used to simulate the shape of the plant. The total area and the number of leaves of each plant will also be calculated. The total area is obtained with the number of the plant pixels and the image calibration result. The number of leaves is derived by detecting the apex of each leaf. It is achieved by using wavelet transform to identify the local maximum of the distance signal between the contour points and the region centroid.

2. Image Preprocessing

2.1. Image Acquisition. At each pot of *Arabidopsis* plant, a small square of checkerboard, with the side of 12 mm, was used to do image correction and calibration. With the alternating dark and bright grids, the corners of the checkerboard pattern present as very strong features and can be detected easily [15]. The top view of the plant was taken by a camera (Canon EOS 500D), in which the image resolution was set to 3456×2304 pixels. The total number of plants is 100. One of the original images is shown in Figure 1(a).

2.2. Image Correction and Calibration. It is inevitable that the image is distorted. It needs to be corrected for further geometrical measurement. The distortion can be expressed as

$$
\begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \tag{1}
$$

where (x_1, x_2, x_3) is the homogeneous coordinate of a point in real world, (x_1', x_2', x_3') is the coordinate in the image. The matrix H is the transform function. It can be computed with four pairs of points. With the result, the inverse matrix H^{-1} can be obtained and the distortion can be eliminated by convolving it with the image.

In this paper, the corners of the checkerboard were detected to compute the transform function. The checkerboard region was first automatically detected by using (2) to recognize the blue border. With blue border region, the checkerboard region was acquired by filling the holes. The Harris corner detection algorithm was then used to detect the corners [16]. Thus the image could be corrected. Figure 1(b) shows the corner detection result and Figure 1(c) shows the image after correction. Since the side of checkerboard is 12 mm, its real size is 144 mm [2]. The real size of each pixel could also be obtained with (3). Consider

$$
B - R > 0,
$$
$$
B - G > 0, \tag{2}
$$
$$
B > 150,
$$

where R, G, and B are the red, green, and blue values of a pixel. Consider

$$
A_{dz} = \frac{144}{N_b}, \tag{3}
$$

where N_b is the number of checkerboard region pixels and A_{dz} is the real size of each pixel.

3. Image Segmentation

The region of green plant in an image was segmented by using Excess Green (ExG) minus Excess Red (ExR) index [16]. It is given as

$$
\text{ExG} - \text{ExR} = 2g - r - b - (1.4r - g), \tag{4}
$$

where r, g, and b are the red, green, and blue chromatic coordinates. For a 24-bit colour image, they are given as

$$
r = \frac{R}{R + G + B},
$$
$$
g = \frac{G}{R + G + B}, \tag{5}
$$
$$
b = \frac{B}{R + G + B},
$$

where R, G, and B are the RGB values of a pixel.

The binary images were obtained with fixing threshold 0. Figure 2(a) shows the segmentation result of the plant. With the segmentation result, the number of plant pixels could be obtained. Combined with the calibration result in preprocessing step, the real area of the plant could be calculated. It is given as

$$
A_p = A_{dz} N, \tag{6}
$$

where N is the number of pixels of the plant region.

The holes inside the region were filled, as shown in Figure 2(b). Thus, the outer overall contour, shown in Figure 2(c), could be obtained.

4. Contour Reconstruction

4.1. Boundary Tracking. After obtaining the plant region, the boundary pixels need to be stored clockwise or counter clockwise. This could be done by using boundary point tracking algorithm [17]. There are three steps to conduct the tracking.

Step 1. Select the pixel of the minimum row coordinate from the pixels of the minimum column value of the region. Let this pixel b_0 as the starting tracking pixel of the boundary. Another pixel c_0 is the west neighbour points of b_0. Store b_0, c_0. Let $b = b_0, \; c = c_0$.

Step 2. Search the 3×3 neighbourhood of b in a clockwise direction from c to find the next boundary pixel b_1. In the 8 neighbour point series of b, the background pixel in front of b_1 is recorded as c_1. Store b_1, c_1. Let $b = b_1, \; c = c_1$.

Step 3. Repeat Step 2 until the pixel b is equal to b_0 and the next boundary pixel is b_1.

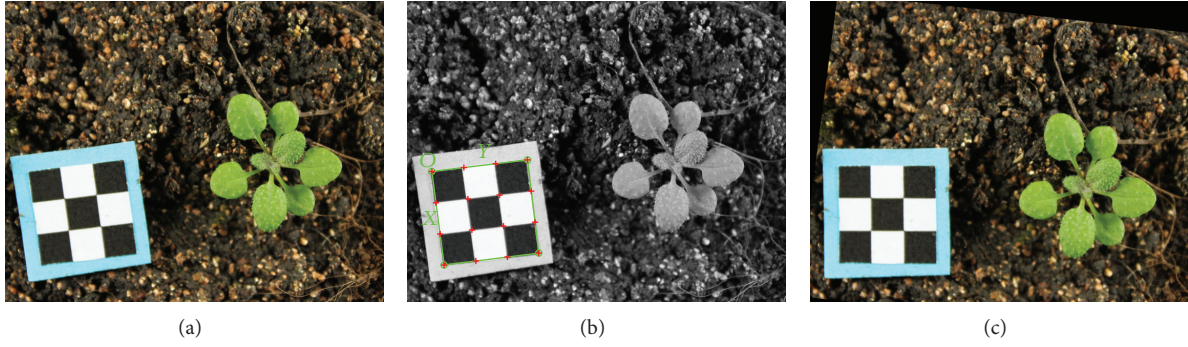With these steps, the boundary point series is obtained.

FIGURE 1: Image preprocessing. (a) Original image of an *Arabidopsis* plant, with soil background. (b) Corner detection result. The corners are marked with red dots. (c) Image correction result.
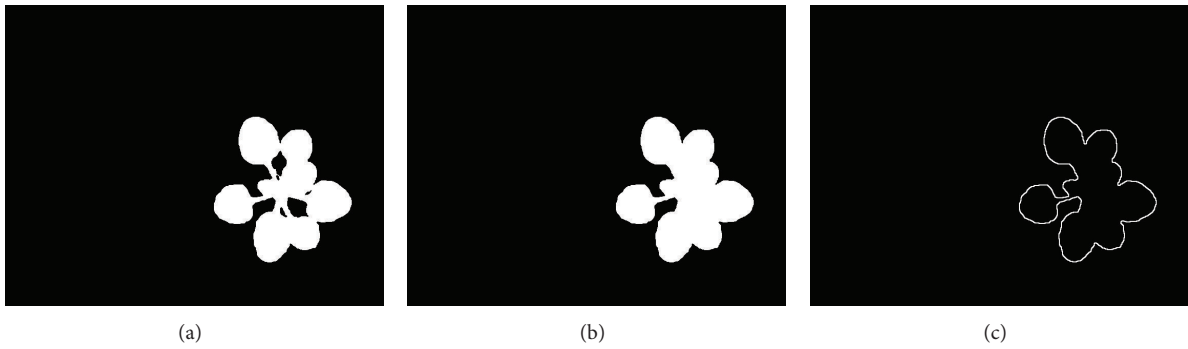

FIGURE 2: Image segmentation. (a) Binary image of segmentation result. (b) Image after filling the holes. (c) Outer contour of the image.

4.2. Elliptic Fourier Descriptors. In the boundary series, each pixel has two values. The image can be viewed as a complex space and the two coordinates can be regarded as the real part and imaginary part. Therefore, the boundary series can be expressed as a discrete function shown in

$$c(i\tau) = x(i\tau) + jy(i\tau) \quad i = 1, 2, \ldots, m, \quad (7)$$

$x(i\tau)$ and $y(i\tau)$ define the two coordinate values of the sampling point i. If the boundary is closed, it can be performed by the *Fourier* expansion in the two coordinate directions to calculate *Fourier* coefficients. It is given as

$$c(i\tau) = \frac{a_{x0}}{2} + \sum_{k=1}^{m/2} \left(a_{xk} \cos(k\omega i\tau) + b_{xk} \sin(k\omega i\tau) \right)$$

$$+ j \left(\frac{a_{y0}}{2} + \sum_{k=1}^{m/2} \left(a_{yk} \cos(k\omega i\tau) + b_{yk} \sin(k\omega i\tau) \right) \right), \quad (8)$$

where

$$a_{xk} = \frac{2}{m} \sum_{i=1}^{m} x_i \cos(k\omega i\tau), \quad b_{xk} = \frac{2}{m} \sum_{i=1}^{m} x_i \sin(k\omega i\tau),$$

$$a_{yk} = \frac{2}{m} \sum_{i=1}^{m} y_i \cos(k\omega i\tau), \quad b_{yk} = \frac{2}{m} \sum_{i=1}^{m} y_i \sin(k\omega i\tau). \quad (9)$$

Thus, each *Fourier* harmonics has one set of coefficients $(a_{xk}, b_{xk}, a_{yk}, b_{yk})$. m is the number of points. T is the period of the boundary. Generally, it is equal to 2π. τ is the sampling period, which is equal to T/m. ω defines the fundamental frequency. It is equal to $2\pi/T$. Thus, $k\omega i\tau$ is equal to $2\pi ki/m$. The harmonics number k defines the level of detail of the characterization [18]. The higher the number is, the more precise descriptions the coefficients provide. However, they also contain much more noise. Generally speaking, the curve can be represented with less than $m/2$ coefficients.

Figure 3 shows reconstructed contour with different harmonics numbers. It can be seen that the contour is reconstructed well with 30 set of coefficients. As the coefficients increase, there are no more new details of the contour but some noise in the image. In other words, 120 coefficients are enough to store the contour of the plant. The data could be used anytime to reconstruct the contour, providing the same representation as the original contour in Figure 2(c). The original contour contains 1163 pixels. Compared with the original number, the coefficients save a lot of storage.

5. Leaf Apex Recognition

From the overall contour, it can be seen that the distance between the apex of each leaf and the region centroid is a local maximum. In this case, we firstly calculated the distance between each boundary point and the centroid. Thus we
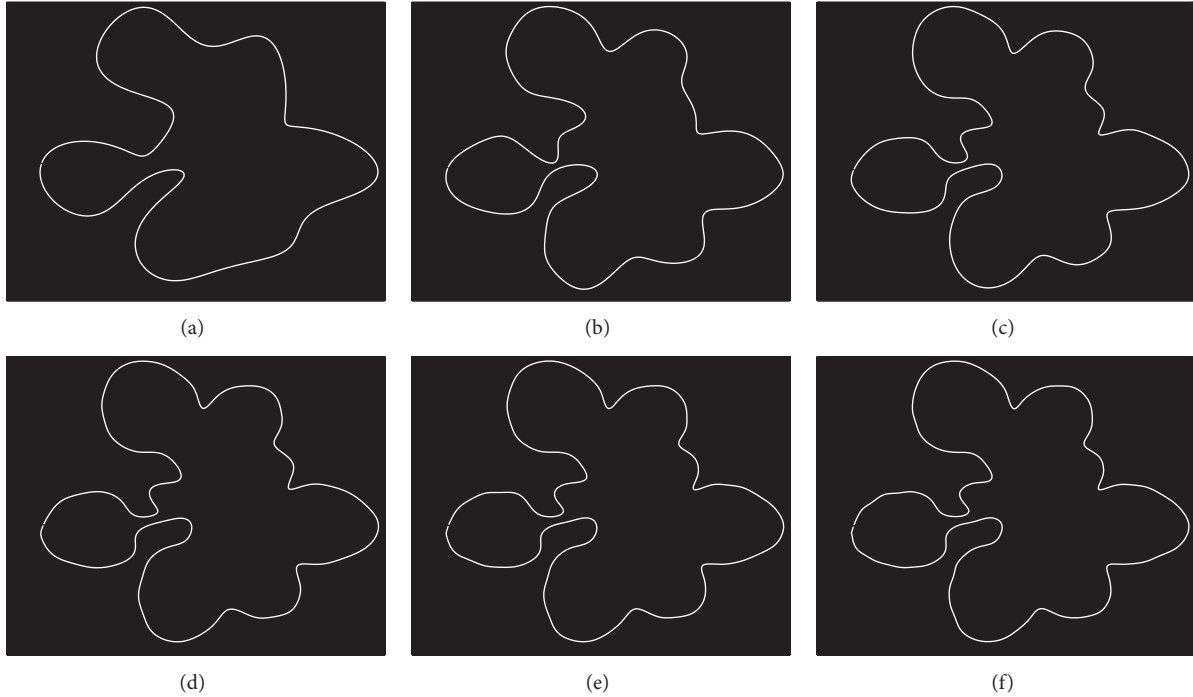
(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 3: Contour reconstruction. (a)–(f) Contour reconstruction with 10, 20, 30, 40, 50, and 60 harmonics numbers, respectively.

obtain a 1D signal. The next step is to identify the local maximum points of the signal.

5.1. 1D Distance Curve. The distance between the contour point and the centroid is calculated as

$$D_i = \sqrt{(x_i - x_o)^2 + (y_i - y_o)^2}, \qquad (10)$$

where (x_i, y_i) is the coordinates of the contour point i, (x_o, y_o) is the values of the centroid point.

5.2. Local Maximum Recognition. The original distance curve signal contains much noise. It is impossible to locate the local maximum only by comparing each contour point with the points at its two sides. In this paper, we used wavelet transform to detect the local maximum point.

Wavelet transform is often used to detect the singularity of the signal. When the signal contains much change, the wavelet coefficients will be quite large. While on the contrary, the wavelet coefficients will be zero if the point is a local minimum or maximum. Another merit of wavelet analysis is that it provides a decomposition of the original signal into different scales. So it is very convenient to choose the scale according to different demands to analyse the signal.

In the experiment, we first expand the signal into 2^n length so that it can be performed by wavelet transform. Then the signal was discomposed into 5 levels with Haar wavelet transform. The wavelet coefficients were used to analyse the local maximum. For a point, if the coefficient of its former one is smaller than 0 while that of its latter one is larger than 0, it will be identified as the zero-crossing point of the wavelet

coefficients, as well as the local maximum of the distance signal. Figure 4(a) shows the wavelet coefficients in the last level. The zero coefficients have been marked with red dots. Figure 4(b) shows the local maximum of 1D distance signal. Figure 4(c) shows the leaf apex detection result.

6. *Arabidopsis* Growth Simulation

The images of *Arabidopsis* plants were taken every day during the growth period. The growth simulation is based on the obtained data in 5 days. The experiments were run on a standard PC with 2.1 GHz Dual-Core CPU and 2 GB memory. The processing time is on average 0.5 second per image using MATLAB. Figure 5 shows one example of the *Arabidopsis* plants. Figures 5(a)–5(e) show the overall contour reconstruction result. They were realized with (8). The harmonics number is 30. The apex of each leaf is also marked in each image. With the leaf apex, the number of leaves can be obtained. Figure 5(f) compares total area of the plant in each day.

7. Conclusion

This paper provides a method to represent the virtual *Arabidopsis* plant at each growth stage. The *Arabidopsis* plant is automatically segmented and the overall contour is obtained. The elliptic Fourier analysis is introduced to extract the coefficients of the contour. With this method, all plant contours could be represented and reconstructed with only 120 coefficients. The data is much less than the original number of contour points, which help to save a lot
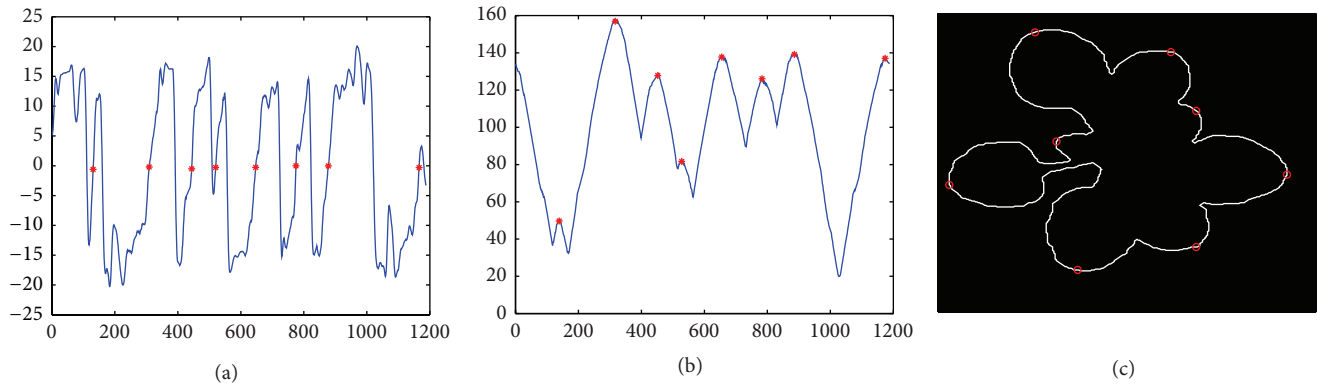
(a)

(b)

(c)

FIGURE 4: Local maximum recognition. (a) Wavelet coefficients in the fifth level. The zero coefficients are marked with red dots. (b) Distance curve between the contour point and the centroid. The local maximum points are labelled with red dots. (c) The apex of each leaf.
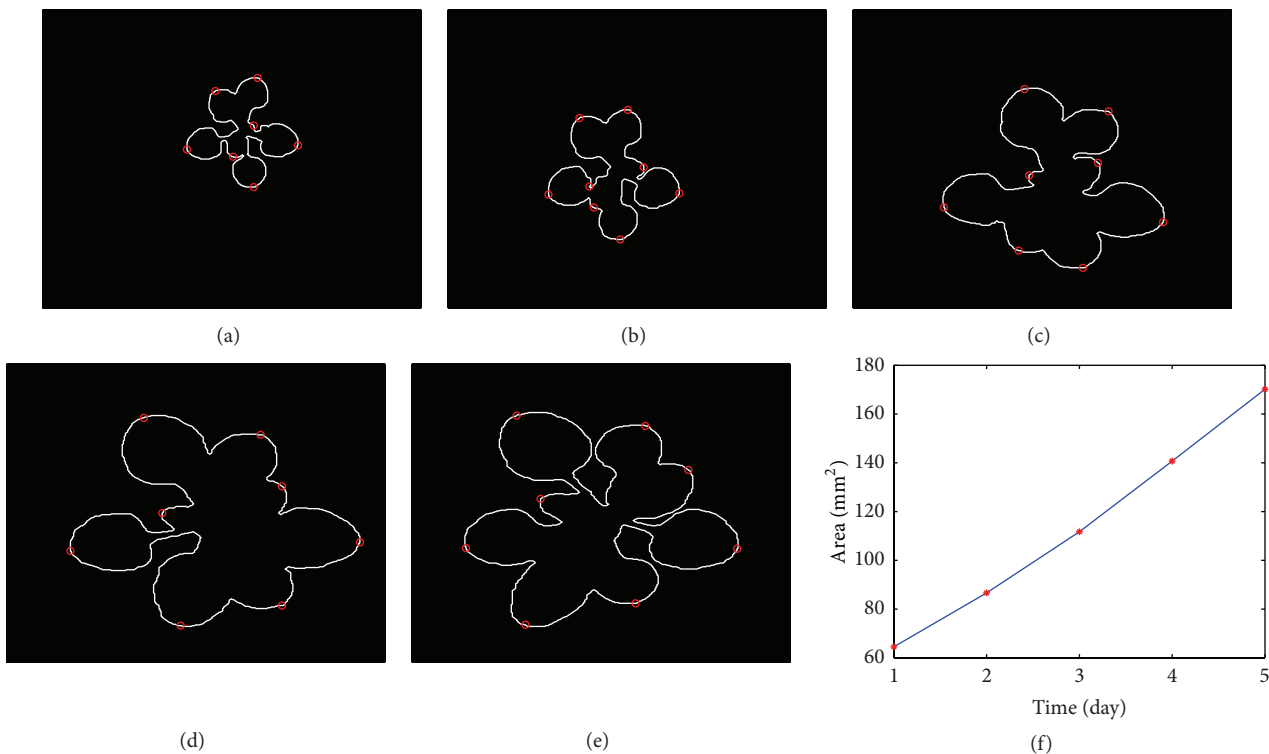

(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 5: Growth simulation. (a)–(e) Contour reconstruction during 5 days. The leaves apexes are marked with red dots. (f) Increase of the total area of the plant.

of storage. In addition, it also makes the data storage much convenient with the same length of coefficients rather than different lengths of contour points. The total area of the plant is measured and the number of leaves is also detected by using wavelet transform to identify the local maximum of the contour-centroid distance signal. Experiment shows that this method can record the growth stage of *Arabidopsis* plant with fewer data and provide a visual platform for plant growth research.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

 [1] P. L. Xu, "Progress of genome study on model plant *Arabidopsis* thaliana," *Shandong Agricultural Science*, vol. 6, pp. 100–102, 2006.

 [2] Arabidopsis Genome Initiative, "Analysis of the genome sequence of the flowering plant *Arabidopsis* thaliana," *Nature*, vol. 408, pp. 796–815, 2000.

 [3] A. Varki and T. K. Altheide, "Comparing the human and chimpanzee genomes: searching for needles in a haystack," *Genome Research*, vol. 15, no. 12, pp. 1746–1758, 2005.

 [4] C. Costa, F. Antonucci, F. Pallottino, J. Aguzzi, D. Sun, and P. Menesatti, "Shape analysis of agricultural products: a review of recent research advances and potential application to computer vision," *Food and Bioprocess Technology*, vol. 4, no. 5, pp. 673–692, 2011.

 [5] F. P. Kuhl and C. R. Giardina, "Elliptic Fourier features of a closed contour," *Computer Graphics and Image Processing*, vol. 18, no. 3, pp. 236–258, 1982.

 [6] J. S. Crampton, "Elliptic Fourier shape analysis of fossil bivalves: some practical considerations," *Lethaia*, vol. 28, no. 2, pp. 179–186, 1995.

 [7] P. E. Lestrel, *Fourier Descriptors and Their Applications in Biology*, Cambridge University Press, Cambridge, UK, 1997.

 [8] A. Loy, S. Busilacchi, C. Costa, L. Ferlin, and S. Cataudella, "Comparing geometric morphometrics and outline fitting methods to monitor fish shape variability of *Diplodus puntazzo* (Teleostea: Sparidae)," *Aquacultural Engineering*, vol. 21, no. 4, pp. 271–283, 2000.

 [9] R. J. Jensen, K. M. Ciofani, and L. C. Miramontes, "Lines, outlines, and landmarks: Morphometric analyses of leaves of Acer rubrum, Acer saccharinum (Aceraceae) and their hybrid," *Taxon*, vol. 51, no. 3, pp. 475–492, 2002.

[10] P. Menesatti, C. Costa, G. Paglia et al., "Shape-based methodology for multivariate discrimination among Italian hazelnut cultivars," *Biosystems Engineering*, vol. 101, no. 4, pp. 417–424, 2008.

[11] C. Costa, P. Menesatti, G. Paglia et al., "Quantitative evaluation of Tarocco sweet orange fruit shape using optoelectronic elliptic Fourier based analysis," *Postharvest Biology and Technology*, vol. 54, no. 1, pp. 38–47, 2009.

[12] N. Furuta, S. Ninomiya, N. Takahashi, H. Ohmori, and Y. Ukai, "Quantitative evaluation of soybean leaf let shape by principal component scores based on elliptic *Fourier* descriptor," *Breeding Science*, vol. 45, pp. 315–320, 1995.

[13] J. Du, D. Huang, X. Wang, and X. Gu, "Shape recognition based on radial basis probabilistic neural network and application to plant species identification," in *Advances in Neural Networks—ISNN 2005*, vol. 3497 of *Lecture Notes in Computer Science*, pp. 281–285, 2005.

[14] J. C. Neto, G. E. Meyer, D. D. Jones, and A. K. Samal, "Plant species identification using Elliptic *Fourier* leaf shape analysis," *Computers and Electronics in Agriculture*, vol. 50, no. 2, pp. 121–134, 2006.

[15] R. C. Gonzalez and R. E. Woods, "Digital image processing," in *Representation and Description*, chapter 11, pp. 511–512,

Publishing House of Electronics Industry, Beijing, China, 3rd edition, 2011.

[16] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proceedings of the 4th Alvey Vision Conference*, pp. 147–151, 1988.

[17] R. C. Gonzalez and R. E. Woods, "Representation and description," in *Digital Image Processing*, pp. 511–512, Publishing House of Electronics Industry, Beijing, China, 3rd edition, 2011.

[18] M. S. Nixon and A. S. Aguado, "Object description," in *Feature Extraction and Image Processing*, chapter 7, pp. 305–307, Academic Press, Oxford, UK, 2008.