SHAPE: A Computer Program Package for Quantitative Evaluation of Biological Shapes Based on Elliptic Fourier Descriptors

H. Iwata and Y. Ukai

Quantitative evaluation of the shapes of biological organs is often required in various research fields, such as agronomy, medicine, genetics, ecology, and taxonomy. Elliptic Fourier descriptors (EFDs), proposed by Kuhl and Giardina (1982), can delineate any type of shape with a closed two-dimensional contour and have been effectively applied to the evaluation of various biological shapes in animals (Bierbaum and Ferson 1986; Diaz et al. 1989; Ferson et al. 1985; Rohlf and Archie 1984) and plants (Furuta et al. 1995; Iwata et al. 1998; McLellan 1993; Ohsawa et al. 1998; White et al. 1988). Quantization of shapes is a prerequisite for evaluating the inheritance of morphological traits in quantitative genetics. There are many reports showing that measurements based on EFDs are helpful for such quantization of the shapes of plant and animal organs. For instance, Iwata et al. (2000) conducted a diallele analysis of the shape of Japanese radish (*Raphanus sativus* L.) roots, using the principal component scores of the EFDs as shape characteristics. Quantitative trait loci (QTL) analysis has also been conducted using the principal component scores of EFDs concerning the shape of the male genitalia of *Drosophila* species (Laurie et al. 1997; Liu et al. 1996). The shape evaluation method based on EFDs can be a powerful tool for analyzing biological shapes, but it is not easy for a researcher to use this method because it involves several complex procedures, such as image processing, contour recording, derivation of the descriptors, and multivariate analysis of the descriptors.

In this article we present SHAPE, a package of programs for evaluating biological contour shapes based on EFDs. This package contains programs for image processing, contour recording, derivation of EFDs, principal component analysis of EFDs, and visualization of shape variations estimated by the principal components. With the aid of this package, a researcher can easily analyze shapes on a personal computer without special knowledge about the procedures related to the method. The principal component scores obtained by the procedures can be used directly as observed values of shape characteristics for the subsequent analyses.

SHAPE is characterized by the following features: (1) The packaged programs are easily operated with the aid of a graphical user interface (GUI); (2) No special computer devices for image processing are required; (3) A large number of samples (say 1,000) can be treated; (4) The scores of principal components are stored in tabbed text format files and can be easily exported for analysis by other software; and (5) The variations in shape accounted for by the principal components can be visualized and printed out.

Quantitative Evaluation of Shape by SHAPE

Image Analysis and Contour Recording

A program named ChainCoder extracts the contours of objects from digital images and stores the relevant information as chain code (Freeman 1975). First, a user takes images of samples using a digital imaging device, such as a digital camera, a photographic film scanner, or an image scanner, and imports them as full-color bitmap files into a computer memory device. ChainCoder takes the images as input files and processes them by the following procedure. The digital image is split into three colors (red, green, and blue [RGB]) with 8-bit quantization (256 gray levels). Then the one image that has the clearest contrast between the objects and the background is converted into a black-and-white image with an appropriate threshold value determined automatically. After noise reduction, the closed contour of the object is extracted by edge detection and the contour information is stored as chain code. The area of the object is also measured and recorded.

Derivation of EFDs

A program named Chc2Nef provides the normalized EFDs from chain-coded contours. The coefficients of EFDs are calculated by the discrete Fourier transformation of a chain-coded contour after Kuhl and Giardina (1982). The coefficients of the EFDs are subsequently normalized to be invariant with respect to the size, rotation, and starting point, with the procedure based on the ellipse of the first harmonic (Kuhl and Giardina 1982). The coefficients can also be normalized with the procedure based on the farthest point on the contour from its centroid. When these methods of normalization are not helpful for giving biological meaning to the contour shape, a user can also normalize the coefficients through manual alignment of the contour. Finally, the normalized coefficients of the EFDs are stored in files.

Principal Component Analysis

A program named PrinComp performs the principal component analysis of the coefficients of the EFDs. The normalized coefficients of the EFDs can still not be used directly as shape characteristics because the number of coefficients is generally very large and the morphological meaning of each coefficient is difficult to interpret separately. Principal component analysis is effective for summarizing the information of the variations contained in the coefficients (Rohlf and Archie 1984). PrinComp performs the principal component analysis based on the variance-covariance matrix of the coefficients. The scores of the derived principal components are also calculated and stored in text format files, which can be provided as input files for the various subsequent analyses, for instance, a quantitative genetic analysis.

Visualization of Shape Variations

A program named PrinPrint visualizes the variation in shape that can be accounted for by each principal component. The visualization is conducted through the procedure proposed by Furuta et al. (1998). First, the coefficients of the EFDs are calculated, letting the score for a particular principal component be equal to the mean plus or minus two times the standard deviation, that is, the square root of the eigenvalue of the component, and the scores of the remaining components be zero. Then the contour shape on each condition can be reconstructed from the calculated coefficients by the inverse Fourier transformation. This visualization may be helpful for giving the morphological meaning of the variation evaluated by each principal component. Reconstructed contours are easily printed on a sheet of paper with an ordinary printer.

Computer Requirements and Availability

The programs contained in SHAPE can be used with any IBM-compatible computer running Windows 95, 98, NT, or higher versions. The program package and a related online manual are available at http://cse.naro.affrc.go.jp/iwatah/shape and http://wheat.ab.a.u-tokyo.ac.jp/~iwata/shape.

From the Department of Information Science and Technology, National Agricultural Research Center, 3-1-1 Kannondai, Tsukuba, Ibaraki 305-8666, Japan (Iwata), and Department of Agricultural Environmental Biology, Graduate School of Agriculture and Life Science, University of Tokyo, 1-1-1 Yayoi, Bunkyo, Tokyo 113-8657, Japan (Ukai). Address correspondence to H. Iwata at the address above, or e-mail: iwatah@affrc.go.jp.

References

Bierbaum RM and Ferson S, 1986. Do symbiotic pea crabs decrease growth rate in mussels? Biol Bull 170:51–61.

Diaz G, Zuccarelli A, Pelligra I, and Ghiani A, 1989. Elliptic Fourier analysis of cell and nuclear shapes. Comp Biomed Res 22:405–414.

Ferson S, Rohlf FJ, and Koehn RK, 1985. Measuring shape variation of two-dimensional outlines. Syst Zool 34:59–68.

Freeman H, 1974. Computer processing of line drawing images. Comp Surv 6:57–97.

Furuta N, Ninomiya S, Takahashi S, Ohmori H, and Ukai Y, 1995. Quantitative evaluation of soybean (*Glycine max* L., Merr.) leaflet shape by principal component scores based on elliptic Fourier descriptor. Breed Sci 45:315–320.

Iwata H, Niikura S, Matsuura S, Takano Y, and Ukai Y, 1998. Evaluation of variation of root shape of Japanese radish (*Raphanus sativus* L.) based on image analysis using elliptic Fourier descriptors. Euphytica 102:143–149.

Iwata H, Niikura S, Matsuura S, Takano Y, and Ukai Y, 2000. Diallel analysis of root shape of Japanese radish (*Raphanus sativus* L.) based on elliptic Fourier descriptors. Breed Sci 50:73–80.

Kuhl FP and Giardina CR, 1982. Elliptic Fourier features of a closed contour. Comp Graphics Image Processing 18:236–258.

Laurie CC, True JR, Liu J, and Mercer JM, 1997. An introgression analysis of quantitative trait loci that contribute to a morphological difference between *Drosophila simulans* and *D. mauritiana.* Genetics 145:339–348.

Liu J, Mercer JM, Stam LF, Gibson G, and Laurie CC, 1996. Genetic analysis of a morphological shape difference in the male genitalia of *Drosophila simulans* and *D. mauritiana.* Genetics 142:1129–1145.

McLellan T, 1993. The roles of heterochrony and heteroblasty in the diversification of leaf shapes in *Begonia dregei* (Begoniaceae). Am J Bot 80:796–804.

Ohsawa R, Tsutsumi T, Uehara H, Namai H, and Ninomiya S, 1998. Quantitative evaluation of common buckwheat (*Fagopyrum esculentum* Moench) kernel shape by elliptic Fourier descriptor. Euphytica 101:175–183.

Rohlf FJ and Archie JW, 1984. A comparison of Fourier methods for the description of wing shape in mosquitoes (*Ritera culicidae*). Syst Zool 33:302–317.

White R, Rentice HC, and Verwist T, 1988. Automated image acquisition and morphometric description. Can J Bot 66:450–459.

Received June 1, 2001
Accepted August 8, 2002

Corresponding Editor: Bruce S. Weir