Identification of *Rhus succedanea* L. Cultivars Using Elliptic Fourier Descriptors Based on Fruit Shape

By Y. HIRAOKA* and N. KURAMOTO

Forest Tree Breeding Center, Kyushu Regional Breeding Office, 2320-5 Suya, Nishigoshi-machi, Kikuchi-gun, Kumamoto 861-1102, Japan

(Received 17th February 2004)

Abstract

We have developed a method to identify cultivars of Rhus succedanea L. based on their fruit contour shape. For this, we collected fruits of five cultivars from three different environments (differing in site and/or year of collection) and the horizontal contour shape of each fruit was expressed by 37 elliptic Fourier descriptors, normalized in terms of size, rotation, shift and starting point of contour tracing. The first six components derived from a principal component analysis of the elliptic Fourier descriptors explained 89% of the variance. The differences among cultivars, environments and the cultivar x environment interaction were significant at the 0.01% probability level for all six principal components according to ANOVA. UPGMA cluster analysis based on the six principal components showed a high degree of clustering and most (but not all) ramets from the same cultivar clustered together. However, results of a UPGMA cluster analysis of Mahalanobis' generalized distances among cultivars and environments, based on the 37 elliptic Fourier descriptors, showed that samples from the same cultivars clustered together, regardless of the environmental factors. We then applied a 'similarity probability' test, based on Mahalanobis' generalized distances and a randomization test. The similarity probabilities between descriptors in the database and sampled fruits, when the cultivars they represented were included in the database, were >97%. In contrast, for samples representing cultivars that were not included in the database, the probabilities were <46%. These figures also apply to pairs of samples included in the database that represented the same cultivar, and different cultivars, respectively. These results suggest that it is possible to identify R. succedanea cultivars based on fruit contour shape using elliptic Fourier descriptors and similarity probability analysis.

Key words: Rhus succedanea L., cultivar identification, fruit shape, elliptic Fourier descriptor, Mahalanobis' generalized distance, randomization test.

Introduction

Rhus succedanea L. (hazenoki in Japanese) is a cultivated tree, the fruit skins of which are used to produce 'haze wax' for high-quality cosmetic items, traditional Japanese candles and other industrial products. Haze wax is a very important product because it contains japanic acid, to which its high viscosity is attributed. *R. succedanea* has been cultivated for over three hundred years and there are many old cultivars in Japan. These cultivars have been propagated by grafting, so each cultivar is clonal. It is clearly important to identify and conserve these cultivars, and to investigate their characters, in order to breed *R. succedanea* successfully using them.

GOTO *et al.* (1997) identified cultivars of R. *succedanea* using RAPD (Random Amplified Polymorphic DNA) markers. DNA markers like RAPDs are useful for identifying plant clones because of their independence from environmental effects.

Silvae Genetica 53, 5-6 (2004)

However, these methods are generally expensive. Thus, if many cultivars have to be identified, cheaper identifying techniques need to be developed.

Traditionally, these cultivars have been identified according to the shape of their organs, e.g. leaves or fruits. Methods of identification based on organ shape have the advantage of being cheaper than techniques based on DNA markers. On the other hand, a major disadvantage of such methods is that organ shape depends not only on genotype, but also on environmental factors.

Thus, for accurate cultivar identification, it is necessary to use traits with relatively low dependence on environmental effects, to characterize many samples and to use approaches that can detect even very slight differences among cultivars. Several previous studies have considered morphological variation among cultivars of R. succedanea. For example, AOKI et al. (1952) identified several typical cultivars of the species using the shapes of their leaves, fruits and crowns, together with their quality and quantity of haze wax. The cited authors concluded that fruit shape was better than other tested traits for identifying cultivars. This group also tried to evaluate minute differences in fruit shape by drawing side-projected figures, then expressing and comparing the resulting shapes for different cultivars in terms of two-dimensional coordinates (AOKI et al., 1953). However, they investigated only three fruits per cultivar, perhaps because of the labor-intensiveness of their approach, and didn't mathematically represent the fruit contour curvature. Thus, differences in fruit shape among cultivars weren't statistically demonstrated. In contrast, recent developments allow the time required to measure the closed contour shape and represent it mathematically to be greatly reduced by using 'the elliptic Fourier method' as processed by image analysis software and a personal computer.

Fourier expansions can be expressed as series of trigonometric functions and they can approximate any periodic function, with no need to formulate any curves like Taylor expansions. Thus, Fourier expansions can be used to approximate unknown curve shapes. Fourier descriptors give good representations of contour shapes, and thus have proved very useful for analyzing such shapes (BIERBAUM and FERSON 1986; BOOKSTEIN et al., 1982; DIAZ et al., 1989; EHRLICH and WEINBERG, 1970; EHRLICH et al., 1983; FERSON et al., 1985; ROHLF and ARCHIE, 1984; WHITE et al., 1988). The elliptic Fourier method is particularly suitable for evaluating closed contour shapes, such as the shapes of leaves or fruits (KUHL and GIARDINA, 1982; ROHLF and ARCHIE, 1984; WHITE et al., 1988). This method has been applied in studies of organs of several plant species, for example, Betula leaf (WHITE et al., 1988), soybean leaflet (FURUTA et al., 1995), Japanese radish root (IWATA et al., 1998; 2000) and citrus leaf (IWATA et al., 2002a; 2002b). In addition, it is possible to summarize the information contained in the elliptic Fourier descriptors by principal component analysis (ROHLF and ARCHIE, 1984), and subsequent application of analysis of

^{*} Corresponding author: YUICHIRO HIRAOKA. Tel: +81-96-242-3151; Fax: +81-96-242-3150. E-mail: <u>yhiraoka@affrc.go.jp</u>

variance to the principal components has detected significant differences among cultivars of some crop species (FURUTA *et al.*, 1995; IWATA *et al.*, 1998), although no attempts were made to identify cultivars in the cited studies.

We believed that analysis of fruit contour shape by elliptic Fourier descriptors would provide a useful method for numerically identifying R. succedanea cultivars. Thus, in the present study we analyzed the fruit contour shape of R. succedanea by this method. Using this approach we also evaluated the variation of fruit shape among cultivars, together with environmental effects, and found that it appears to provide a viable way to identify cultivars that is affected by environmental factors less than some alternative approaches.

Materials and Methods

We selected five traditional cultivars of R. succedanea for this study: Budo, Ikichi, Joh, Oh and Showafuku, and three different ramets per cultivar from each of two sites, Kurume city (in both 2001 and 2002) and Minamata city (in 2001) on Kyushu Island, Japan (*Fig. 1*). In each case, ninety fruits per ramet were randomly sampled from their respective crowns. The samples from these five cultivars were used to investigate genotype and environmental effects on variation of fruit shape, and the results were used to provide a database for cultivar identification.

In order to test the feasibility of using this database to identify cultivars, we sampled a further thirty fruits per cultivar from each of these five cultivars and another two, named Matsuyama and Taikobo.

Sample fruits were each placed on a scanner (GT-800, SEIKO EPSON Corp., Nagano) horizontally and scanned images of the fruits were stored in a personal computer. The binary image of each fruit was obtained by thresholding the original red digital image of the fruit. After noise reduction, the closed contour of each fruit was extracted by edging the binary image and the extracted contour was then chain-coded (FREE-MAN, 1974).

Each chain-coded contour was projected onto the *XY* plane so that the starting point of the chain-code was located at the origin of the coordinate. Assuming that a mass point moves along



Figure 1. – The map of sampling sites.

the contour at the same speed, the x and y coordinates of the mass point periodically oscillate and can be described by a Fourier series expansion. This expression for a closed contour is called the elliptic Fourier descriptor. If the contour between the (*i*-1)-th and the *i*-th chain-coded points is linearly interpolated, and the length of the contour from the starting point to the *p*-th point and the perimeter of the contour are denoted t_p and T, respectively, then,

$$t_P = \sum_{i=1}^P \Delta t_i$$

and $T = t_k$, where Δt_i and K are the distance between the (i-1)th and the *i*-th points and the total number of the chain-coded points on the contour, respectively. Note that the *K*-th point is equivalent to the starting point. Let the *x* coordinate of the *p*-th points be denoted x_p . Then,

$$x_p = \sum_{i=1}^p \Delta x_i ,$$

where Δx_i is the distance along the *x* axis between the (*i*-1)-th and the *i*-th point. The elliptic Fourier expansions of the coordinates on the contour are

$$x_p = A_0 + \sum_{n=1}^{\infty} \left(a_n \cos \frac{2n\pi t_p}{T} + b_n \sin \frac{2n\pi t_p}{T} \right)$$

Then, assuming linear interpolation between the neighboring points, the elliptic Fourier coefficients of the *n*-th harmonic (a_n, b_n) are given as

$$a_{n} = \frac{T}{2n^{2}\pi^{2}} \sum_{p=1}^{K} \frac{\Delta x_{p}}{\Delta t_{p}} (\cos \frac{2n\pi t_{p}}{T} + \cos \frac{2n\pi t_{p-1}}{T})$$

and

$$b_n = \frac{T}{2n^2 \pi^2} \sum_{p=1}^{K} \frac{\Delta x_p}{\Delta t_p} (\sin \frac{2n\pi t_p}{T} + \sin \frac{2n\pi t_{p-1}}{T}).$$

We also calculated $\boldsymbol{c}_{\boldsymbol{n}}$ and $\boldsymbol{d}_{\boldsymbol{n}}$ (corresponding coefficients for the y coordinate).

In this study, we used the first ten harmonics to express the contour shape of the fruit of *R. succedanea*. Because these coefficients are not constant in size, rotation, shift or starting point of chain-coding for a contour, they cannot be used as indicators to compare the shape of objects. For example, the Fourier coefficients differ even for exactly the same shape when the starting point of the chain-coding differs. So, in this study, the Fourier coefficients were normalized according to KUHL and GIARDINA (1982) (see Appendix). This normalizing procedure is based on the ellipse of the first harmonic. Since normalized a_1 , b_1 and c_1 become constant, we then have 37 normalized coefficients excluding the constant coefficients ($[10 \times 4] - 3$).

We then subjected the data to PCA (principal component analysis) in order to summarize the information contained in the elliptic Fourier descriptors and ANOVA (analysis of variance) of the principal components in order to evaluate statistically the effects of differences among 'cultivar' and environmental factors, and their interaction, on fruit shape. In addition, we carried out UPGMA (unweighted pair group) cluster analysis to evaluate visual differences in fruit shape among cultivars and/or environments. We also calculated Mahalanobis' generalized distances. Whereas the basic Euclidian distance treats each variable as equally important in calculating the distance, this method scales the contribution of individual variables to the distance value according to the variability of each variable. This distance is a measure of the distance between each observation in a multidimensional cloud of points and the centroid of the cloud. The Mahalanobis' generalized distance D^2 is given by

$$D^2 = (\mathbf{x} - \mathbf{m})^{\prime} \mathbf{V}^{-1} (\mathbf{x} - \mathbf{m})$$

where \mathbf{x} is a vector of values for a particular observation, \mathbf{m} is the vector of means of each variable, and \mathbf{V} is the variancecovariance matrix.

We also performed a randomization test to test the significance of cultivar identifications. For this, 30-fruit datasets (each including 37 variables) were selected randomly from the database, avoiding overlap, until every fruit's descriptors were included in a dataset. Means were then calculated for each of the elliptic Fourier descriptors. Thus, since there were 1,350 datasets (five cultivars x 270 fruits), we obtained 45 mean sets of each descriptor per trial. This calculation was repeated 100 times, giving 4,500 mean sets. Then, we calculated Mahalanobis' generalized distances between these 4,500 datasets and the five cultivars' datasets in the database and generated a frequency distribution of the distances (giving 22,500 datasets, i.e. 4,500 x 5). After that, we also calculated Mahalanobis' generalized distances between each mean of the 30 samples selected for the identification test and each cultivar mean in the database, and derived the probability of this distance occurring in the frequency distribution (the 'similarity probability'). The resulting probability represents the relative closeness between each sample dataset and each dataset included in the database in 37-dimensional space.

The elliptic Fourier descriptors and the PCA were calculated with SHAPE Ver. 1.2 (Iwata, 2001), a free software package, and the other analyses with R: a computer programming language for statistical analysis.

Results

Differences in fruit shape among cultivars and environments represented by the principal component scores

Table 1 shows the results of the PCA based on the elliptic Fourier descriptors. The contributions of the first, second and third principal components were 57%, 13% and 7%, respectively, and the first six principal components cumulatively explained 89% of the variance. *Fig.* 2 shows the estimated fruit shape obtained by redrawing the contour with the principal components. The first, second and third principal components were related to the length to width ratio, the location of the protrusion and the location of the centroid, respectively. The fourth to sixth principal components related to variations in contour shape that were difficult to describe non-mathematically.

We carried out two-way ANOVA to investigate the effects of cultivar (five variations) and environment (three variations: sampled at Kurume in 2001 or 2002, or at Minamata in 2001) for the six principal components (*Table 2*). The differences among cultivars and environments and the cultivar \times environment interaction were found to be significant at the 0.01% probability level for all six principal components. The contributions of the 'cultivar' factor were relatively high in the first, second and fourth principal components. In addition, we

Table 1. - Summary of the principal component analysis.

Principal	Eigenvalue	Contribution	Cumulative
component	(10^{-3})	(%)	(%)
PC1	1.0519	56.9	56.9
PC2	0.2396	13.0	69.8
PC3	0.1267	6.8	76.7
PC4	0.0996	5.4	82.1
PC5	0.0713	3.9	85.9
PC6	0.0591	3.2	89.1



Figure 2. – Contour shape of the fruit drawn using the elliptic Fourier descriptors estimated under six principal component scores (PC). Each column shows that the score takes either +2S.D., mean or -2S.D. The leftmost column shows the overlaid drawings of the three cases.

carried out a UPGMA cluster analysis of the six principal components (*Fig. 3*), finding that the cultivars generally formed separate clusters, except that "Ikichi B" clustered with "Joh".

Application of 'similarity probabilities'

Another UPGMA cluster analysis was carried out based on Mahalanobis' generalized distances among means of all the elliptic Fourier coefficients for each ramet (*Fig. 4*). Ramets from the same cultivar invariably clustered together. There was no relationship between similarity of fruit shape and the environmental factors, i.e. sampling year or site.

Table 3 shows the 'similarity probabilities' between seven sample sets and five cultivars in the database. The probabilities between the same cultivars were more than 97% in all cases. In contrast, the probabilities between different cultivars were all low: less than 46%. For "Matsuyama" and "Taikobo", which were not included in the database, the probabilities were also low. These findings suggest that it is possible to identify *R. succedanea* cultivars using this method.

Discussion

Merits of the elliptic Fourier method for contour shape analysis

IWATA et al. (1998) listed several advantages of the elliptic Fourier method: it requires neither landmarks nor prior knowl-

Table 2. - F-value and contribution of each factor of the ANOVA.



Figure 3. – The dendrogram based on six principal components of the fruit shape. The alphabet followed each cultivar name indicates the environments. A: fruits sampled at Kurume city in 2001; B: sampled at Kurume city in 2002; C: sampled at Minamata city in 2001.

edge about the shape variation of the target objects; it can visualize the contour information and reconstruct the original shape; it can be mathematically normalized to size, rotation and starting point of the contour trace; and it can be performed automatically using computer programs. In the present study, we confirmed some of these merits when using the elliptic Fourier method to analyze *R. succedanea* fruit shapes.

We were able to express variations in R. succedanea fruit shape, ranging from major to minor, using elliptic Fourier descriptors. AOKI *et al.* (1953) tried to evaluate differences in fruit shape amongst cultivars of R. succedanea by drawing side-projected figures and expressing them in terms of twodimensional coordinates. However, whereas they focused on parts of the fruit contour curvature that clearly differed among cultivars, we treated the whole fruit contour using the method described above. Furthermore, it was possible to visualize fruit shape variations, that were difficult to describe verbally, using the results of the PCA. These findings suggest that the elliptic Fourier method is especially useful for the quantitative measurement of contour shapes of objects with few landmarks in their outline, like the fruits of R. succedanea.

We were also able to avoid the laboriousness of previous methodologies. In order to identify cultivars from the shape of plant organs, it is necessary to use many samples. The method of AOKI *et al.* (1953) is very laborious for describing the shape of a single fruit, so it is difficult to measure large numbers of

	_	PC1		PC2		PC3		PC4		PC5		PC6	
Factor	DF	F	Contribution	F	Contribution	F	Contribution	F	Contribution	F	Contribution	F	Contribution
Cultivar	4	251.0	0.803	338.4	0.474	15.5	0.037	284.2	0.432	46.0	0.106	57.4	0.122
Environment	2	17.2	0.027	19.1	0.013	21.9	0.027	13.4	0.009	67.9	0.079	81.4	0.087
Cultivar × Environment	8	13.6	0.087	15.5	0.041	13.9	0.066	15.5	0.044	5.4	0.021	16.3	0.066
Error			0.083		0.473		0.869		0.514		0.794		0.726

Mahalanobis' generalized distance



Figure 4. – The dendrogram based on the elliptic Fourier descriptors of the fruit shape using Mahalanobis' generalized distance. Data names are described in the similar way as in Fig. 3.

 $Table \ 3.$ – The similarity probabilities (%) of sample cultivars on the database.

_	Cultivar in the database						
Sample cultivar	Budo	Ikichi	Joh	Oh	Showafuku		
Budo	97.42	0.30	3.34	0.86	0.03		
Ikichi	2.62	99.80	33.46	11.56	1.22		
Joh	29.38	2.14	99.70	1.74	0.09		
Oh	0.05	3.10	14.11	99.74	0.89		
Showafuku	4.15	18.90	15.90	6.28	99.58		
Matsuyama	0.03	6.27	0.20	46.24	0.54		
Taikobo	2.10	37.62	24.91	0.31	38.23		

samples by this means. This problem can be solved by automatic measurement of digitized fruit profiles using a personal computer, an image scanner and image analyzing software. In this study, 90 fruits per ramet and 270 per cultivar were sampled to calculate elliptic Fourier descriptors, which appeared to give sufficient data to assess the effects of 'cultivar' and environmental effects on variations in fruit shape. AOKI *et al.* (1953) manually investigated only three fruits per cultivar, perhaps because of the time requirements for their method. In contrast, in the present study almost all measurements were automatically done by computer with software for analyzing closed contour shapes based on elliptic Fourier descriptors.

In conclusion, the elliptic Fourier method used in this study provides a convenient and powerful tool for measuring R. succedanea fruit shapes.

Identification of R. succedanea cultivars using the elliptic Fourier descriptors

According to the ANOVA, each principal component that was related to fruit shape was affected by both 'cultivar' and the environmental factors, and the contribution of the 'cultivar' factor was relatively high for each component. In addition, almost all ramets of the same cultivars clustered in the cluster analysis. These results suggest that the fruit shape of *R. succedanea* is influenced more by genetic factors than environmental factors (although it is strongly affected by both types of factor).

The cluster analysis helped elucidate the relationships among the cultivars' fruit shapes. While it was difficult to group samples according to cultivar by UPGMA cluster analysis based on principal components, they could be classified using Mahalanobis' generalized distances of all the elliptic Fourier descriptors based on fruit shape. These results suggest that Mahalanobis' distances for all the coefficients obtained from the analysis could be successfully used to represent differences among cultivars. Nevertheless, it is difficult to determine whether sample fruit sets were from the same cultivar or not using the results of either the ANOVA or the cluster analysis because of the lack of criteria for such classification, and it might be difficult to assign a new and unknown sample to any cultivar using the cluster analysis.

In addition, we tried to classify cultivars using discriminant analysis based on all descriptors, but this method was also not very successful. Even if two samples were from the same cultivar, they tended to be discriminated as significantly different groups, because the descriptors were strongly affected by environmental factors.

In contrast, the 'similarity probability' approach introduced here was able to assign new samples to cultivars represented in the database, or give indications that they represented cultivars that were not included in it. Using the randomization test, Mahalanobis' distances and these probabilities we were able to correctly identify all the sample cultivars included in the database. The success of this approach was attributed to the facts that it represents the relative closeness between samples and cultivars included in the database in high-dimensional space and that the genetic variance was greater than the environmental variance. However, it will be necessary to add further cultivars to generate a database that allows more cultivars to be identified.

Requiring only an image scanner and a personal computer, this method is substantially cheaper than DNA analysis and no special skills are required. In addition, it enabled a sample set of fruits from an unidentified cultivar to be classified without destroying the fruits, thus allowing them to be used for other purposes later. DNA analysis would be destructive in such cases. This is another advantage of the method presented here over DNA analysis. Further, it would be possible to screen large samples without destroying them using this method, and to test their homogeneity or heterogeneity (allowing mixed fruit lots to be sorted according to their respective cultivars).

In summary, we were able to identify *R. succedanea* cultivars based on fruit shape as expressed by the elliptic Fourier descriptors and similarity probabilities without DNA analysis. This method could be used in the management of old cultivars or plus trees, and in *R. succedanea* breeding programs.

Acknowledgments

We are grateful to Mr. M. OGATA for his maintenance of the study sites. We thank Dr. K. IKEDA and the staff at Fukuoka Prefecture Forest Reserve & Extension Center for collecting sample fruits. We also thank Prof. S. SHIRAISHI, Dr. H. IWATA and Dr. Y. FUJISAWA for their helpful advice.

References

AOKI, Y., YAMAUCHI, M. and NAKASHIMA, Y. (1952): Study of identification in *Rhus succedanea* L. Fukuoka Pref. Forest Exp. Stn.: 1–85 (in Japanese).

- AOKI, Y., YAMAUCHI, M. and TAKESHITA, K. (1953): Study of identification in *Rhus succedanea* L. (2) Identification based on fruit shape. Fukuoka Pref. Forest Exp. Stn.: 1–25 (in Japanese).
- BIERBAUM, R. M. and FERSON, S. (1986): Do symbiotic pea crabs decrease growth rate in mussels? Biol. Bull. **170**: 51–61.
- BOOKSTEIN, F. L., STRAUSS, R. E., HUMPHRIES, J. M., CHERONOFF, B., ELDER, R. L. and SMITH, G. R. (1982): <u>A</u> comment upon the uses of Fourier methods in systematics. Syst. Zool. **31**: 85–92.
- DIAZ, G., ZUCCARELLI, A., PELLIGRA, I. and GHIANI, A. (1989): Elliptic Fourier analysis of cell and nuclear shapes. Comp. Biomed. Res. **22**: 405–414.
- EHRLICH, R. and WEINBERG, B. (1970): An exact method for characterization of grain shape. J. Sed. Pet. **40**: 205–212.
- EHRLICH, R. B., PHARR, JR. and HEALY-WILLIAMS, N. (1983): Comments on the validity of Fourier descriptors in systematics: <u>A Reply to BOOKSTEIN *et al.*</u> Syst. Zool. **32**: 202–206.
- FERSON, S., ROHLF, F. J. and KOEHN, R. K. (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, N., 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.
- GOTO, S., WATANABE, A. and IKEDA, K. (1997): Use of RAPD markers for cultivar identification in *Rhus succedanea* L. J. Jpn. For. Soc. **79**: 229–233 (in Japanese with English summary).
- 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.
- IWATA, H. (2001): SHAPE Ver. 1.2. A software package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors. Free program distributed by the author over the internet from the Home Page at http://cse.naro.affrc.go.jp/iwatah/shape/index.html.
- IWATA, H., NESUMI, H., NINOMIYA, S., TAKANO, Y. and UKAI, Y. (2002a): Diallel analysis of leaf shape variations of citrus varieties based on elliptic Fourier descriptors. Breed. Sci. **52**: 89–94.
- IWATA, H., NESUMI, H., NINOMIYA, S., TAKANO, Y. and UKAI, Y. (2002b): The evaluation of genotype environment interactions of citrus leaf morphology using image analysis and elliptic Fourier descriptors. Breed. Sci. 52: 243–252.
- KUHL, F. P. and GIARDINA, C. R (1982): Elliptic Fourier features of a closed contour. Comp. Graph. Image Proc. 18: 236–258.
- ROHLF, F. J. and ARCHIE, J. W. (1984): A comparison of Fourier methods for the description of wing shape in mosquitoes (*Ritera culicidae*). Syst. Zool. **33**: 302–317.
- WHITE, R. J., RENTICE, H. C. and VERWIST, T. (1988): Automated image acquisition and morphometric description. Can. J. Bot. **66**: 450–459.

Appendix

The procedure for standardizing the Fourier coefficients, according to Kuhl and GIARDINA (1982), is as follows. Let the standardized coefficients of the n-th harmonic be a_n^{**} , b_n^{**} , c_n^{**} and d_n^{**} . Then,

$$\begin{bmatrix} a_n^{**} & b_n^{**} \\ c_n^{**} & d_n^{**} \end{bmatrix} = \frac{1}{E^*} \begin{bmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{bmatrix} \begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix} \begin{bmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{bmatrix}$$

where

$$E^* = (a_1^{*2} + c_1^{*2})^{1/2},$$
$$\begin{bmatrix}a_1^*\\c_1^*\end{bmatrix} = \begin{bmatrix}a_1 & b_1\\c_1 & d_1\end{bmatrix} \begin{bmatrix}\cos\theta\\\sin\theta\end{bmatrix}$$
$$\psi = \arctan\left[\frac{c_1^*}{a_1^*}\right]_{(0 \le \psi < 2\pi)}$$

and

$$\theta = \frac{1}{2} \arctan\left[\frac{2(a_{1}b_{1} + c_{1}d_{1})}{(a_{1}^{2} + c_{1}^{2} - b_{1}^{2} - d_{1}^{2})}\right](0 \le \theta < \pi)$$

This standardizing procedure is based on the ellipse of the first harmonic. E^* and ψ are the length and direction of the major axis of the first harmonic ellipse, respectively. These are parameters for the size-invariance and rotation invariance, respectively. Because there are two solutions of ψ , we choose one of them, based on the position of the pulvinus of each sample fruit relative to the first harmonic ellipse. θ is the phase shift of the radius vector of the first harmonic ellipse at p = 0 and the parameter for the chain-code starting point invariance. The shift-invariance is obtained simply by setting the location parameters of the Fourier expansions (A_0 and C_0) to zero. This standardization makes a_n^{**} , b_n^{**} , c_n^{**} and d_n^{**} independent of the size, rotation, shift and chain-code starting point of a contour.