AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY 86:415–427 (1991)

Euclidean Distance Matrix Analysis: A Coordinate-Free Approach for Comparing Biological Shapes Using Landmark Data

SUBHASH LELE AND JOAN T. RICHTSMEIER
Department of Biostatistics, The Johns Hopkins University, School of Hygiene and Public Health (S.L.) and Department of Cell Biology and Anatomy, The Johns Hopkins University, School of Medicine (J.T.R.), Baltimore, Maryland 21205

KEY WORDS Bootstrap, *Cebus apella*, Euclidean distance matrix, Sexual dimorphism, Morphometrics, Form analysis, Union–intersection test

ABSTRACT For problems of classification and comparison in biological research, the primary focus is on the similarity of forms. A biological form can be conveniently defined as consisting of size and shape. Several approaches for comparing biological shapes using landmark data are available. Lele (1991a) critically discusses these approaches and proposes a new method based on the Euclidean distance matrix representation of the form of an object. The purpose of this paper is to extend this new methodology to the comparison of groups of objects. We develop the statistical versions of various concepts introduced by Lele (1991a) and use them for developing statistical procedures for testing the hypothesis of shape difference between biological forms. We illustrate the use of this method by studying morphological differences between normal children and those affected with Crouzon and Apert syndromes and craniofacial sexual dimorphism in *Cebus apella*.

To be able to quantitatively compare the shapes of biological objects, we need a method for cataloguing the forms under consideration. This can be done in several ways. The choice of the type of data used in analysis depends upon the nature of the biological object under study, as well as the focus of the investigation. We feel that when available, landmark data have certain advantages over other mensurable components. The most important advantage is maintenance of the relative position of all biological loci of interest, or of the geometric integrity of the form as represented by landmarks. Only specially designed suites of linear measurements based on landmark data maintain the geometric integrity of the forms under study. Series of measurements made up of external dimensions such as maximum breadth of a structure, or minimum diameter of a feature are inappropriate for the methods introduced here. In this study we use biological landmark coordinate data to archive each of the forms used in analysis. Since certain biological implications are involved in the use of landmark data, we discuss them before presenting the method.

LANDMARK DATA

Most biological forms contain specific loci referred to as biological landmarks. Landmarks are structurally consistent loci which can have evolutionary, ontogenetic, and/or functional significance, and must be consistently present on all forms under consideration in order to be useful in analysis. Landmarks are often referred to as "homologous" points. Homology is used here in the sense given by Van Valen (1982) and further discussed by Roth (1988), as a correspondence between two or more characteristics of organisms that is caused by continuity of information. The minimal criterion for a feature, character, or landmark to be used as a homologous point in morphometric analysis is that given a single definition, it can be consistently and reliably located with a mensurable degree of accuracy on all forms considered.

When archiving landmark locations of a three dimensional form, a $K \times 3$ matrix is produced where K is the number of land-

Received December 28, 1989; accepted March 12, 1991.

marks on that form. Data appropriate for analysis by methods proposed in this paper include $K \times 2$ or $K \times 3$ matrices of coordinates. The number of landmarks, K, depends largely on the nature of the forms under study or the research question being addressed. In a complex form like the mammalian skull, there are a large number of landmarks which can be used as homologous points. Because the neurocranium is made up of relatively large, smooth bones with fewer sutural intersections, foramina, and bony prominences, description of the neurocranium by landmark coordinate data is less thorough than for the face, for example. Analytical results are consequently less complete for the neurocranium.

When data are collected from images of biological forms (e.g., X-rays, computed tomography scans, positron emission tomography scans) landmark identification can become more difficult. The identification of landmarks on images is dependent upon certain characteristics of the images. The result may be fewer landmarks, or landmarks of a different kind. For these reasons landmark data collected from different sources are often not directly comparable.

There are biological forms on which very few, or no landmarks exist, or on which landmarks are not easily located. Alternative statistical techniques for the comparison of forms using alternate data (e.g., Fourier series approximations) are appropriate in those cases. We will not discuss such data or methods here.

Several morphometric methods have been proposed for the analysis of landmark data, each with merits and demerits. In Lele (1991a) problematic issues associated with superimposition methods (Bookstein, 1986; Goodall and Bose, 1987; Goodall, 1991) and finite element scaling analysis (Lewis et al., 1980; Cheverud and Richtsmeier, 1986) are discussed. Bookstein's (1989) thin plate splining methods also contain an element of subjectivity since the choice of spline function is based upon mathematical properties rather than a biological model. Lele (1991a) suggests an alternative approach based on the Euclidean distance matrix representation of the objects under study that overcomes the problems associated with other methods and is justified on various biological and statistical grounds.

Lele and Richtsmeier (1990) have shown that statistical models used in morphometric analysis are often inappropriate for biological and statistical reasons. The development of methods for the statistical analysis of landmark data that are more flexible and less model dependent than the existing ones is clearly needed. This paper presents our preliminary attempt toward that goal.

The purpose of this paper is to present statistical versions of various concepts described in Lele (1991a) and to develop a testing procedure for shape differences. In addition, we illustrate our approach using two examples: (1) a comparison of craniofacial morphology of normal children with that of children with Crouzon and Apert syndrome in two dimensions; and (2) a comparison of male and female facial morphology in *Cebus apella* in three dimensions. A methodology for localizing morphological differences with application to similar data sets is discussed in the companion paper (Lele and Richtsmeier, 1991). Throughout this paper, we use the notation developed in Lele (1991a).

SOME DEFINITIONS

We assume that homologous landmarks occur on every object to be compared. The coordinates of these landmarks serve as raw data. Let there be K landmarks and P dimensions. Usually P will be equal to two or three.

Let $\mathbf{X}$ be a matrix of landmark coordinates with K rows and P columns: the i^{th} row consists of the P coordinates of the i^{th} landmark. Note that given $\mathbf{X}$ one can approximate the relative location of the landmarks of the object. Let $\mathbf{F(X)}$ denote the Euclidean distance matrix, henceforth referred to as the *form matrix* (see Lele, 1991a) corresponding to the object with landmark coordinate matrix $\mathbf{X}$. $\mathbf{F(X)}$ is a symmetric matrix of dimension $K \times K$ that consists of distances between all possible pairs of landmarks.

We define various quantities in terms of $\mathbf{X}$ as well as $\mathbf{F(X)}$. Note that Bookstein (1986) and Goodall and Bose (1987) describe their statistical model solely in terms of $\mathbf{X}$—the landmark coordinate matrix.

Below we generalize definitions concerning equality of two forms to equality of two populations of forms. This can be done in several ways. We have chosen to generalize these definitions in two ways: (1) in terms of the identical distribution of forms, and (2) in terms of equality of mean forms.

We assume that $\mathbf{X}$ is some matrix valued random variable. (From this point on, we use the term random variable as a shortened form of matrix valued random variable.) For

example, under the Gaussian perturbation model used by Bookstein (1986) and Goodall and Bose (1987), **X** has a matrix valued Gaussian distribution. We now define equality of form and equality of shape in terms of the matrix valued random variables **X** and **Y**, where **Y** is a matrix of landmark coordinates for form Y. (**X** and **Y** are always matrix valued.)

Definition 1

Two random variables **X** and **Y** are said to have the same form if after proper rotation and translation **X** and **Y** are *identically distributed*. That is: $\mathbf{X} \stackrel{d}{=} \mathbf{YB} + 1_k t^T$, for some orthogonal matrix **B** and a vector t. By identically distributed we mean that the probability distribution functions for **X** and **Y** are the same, although particular observations could and would be different.

Definition 2

Two random variables **X** and **Y** are said to have the same shape if after proper translation, rotation and scaling **X** and **Y** are *identically distributed*. That is: $\mathbf{X} \stackrel{d}{=} b\mathbf{YB} + 1_k t^T$, for some scalar $b > 0$, **B** and t as above.

The corrsponding definitions in terms of the form matrix are:

Definition 3

Two random variables **X** and **Y** are said to have the same shape if $\mathbf{F(X)} \stackrel{d}{=} c\mathbf{F(Y)}$ for some scalar $c > 0$. If $c = 1$, then they have the same form.

In practice it is difficult to test the hypotheses of equality of two distributions ($\mathbf{X} \stackrel{d}{=} \mathbf{Y}$), especially for matrix valued random variables. This is one of the problems that occurs for trivariate and higher dimensional random variables. The data are too sparse to conduct nonparametric tests. We give simplified versions of equality of form and shape below in terms of mean form and mean shape. Let $E(\cdot)$ denote the expectation operator. For example, **E(X)** denotes the average of the random variable **X**, or the average form representing a sample of forms.

Definition 4

We say that random matrices X and Y are equal in *mean form* if and only if

$$\mathbf{E(X)} = \mathbf{E(Y)}\,\mathbf{B} + 1_K t^T$$

for some orthogonal matrix **B** and a vector t, i.e., after translation and rotation the means of **X** and **Y** are equal.

Definition 5

We say that random matrices X and Y are equal in *mean shape* if and only if

$$\mathbf{E(X)} = b\mathbf{E(Y)}\,\mathbf{B} + 1_K t^T$$

for some scalar $b > 0$, **B** and t as above, i.e., after translation, rotation, and scaling, the means of **X** and **Y** are equal.

Note: The independent observations $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ (i.e., the landmark coordinate matrices) from the distribution of X are *identically distributed* only after proper translation and rotation. The same caution applies to $\mathbf{Y}_i$'s, the observations from the distribution of **Y**. In this paper we are dealing with form matrices $\mathbf{F}(\mathbf{X}_i)$ and $\mathbf{F}(\mathbf{Y}_i)$ which are invariant under rotation and translation and are therefore identically distributed.

We now give the same definitions in terms of form matrices.

Definition 6

Given two random variables **X** and **Y** we say that they are equal in *mean shape* if and only if $\mathbf{F[E(X)]} = c\mathbf{F[E(Y)]}$ for some scalar $c > 0$. By this we mean that two mean forms have the same shape if one form is a scaled version of the other. If $c = 1$ then they are equal in mean shape.

To examine the differences between two average forms we propose the use of a matrix of ratios of corresponding linear distances measured on **X** and **Y**. We call this matrix the average form difference matrix.

Definition 7

Given two random variables **X** and **Y**, we define the average form difference matrix by

$$\mathbf{D[E(X), E(Y)]} = [D_{ij}[\mathbf{E(X)}, \mathbf{E(Y)}]] = F_{ij}[\mathbf{E(X)}] / F_{ij}[\mathbf{E(Y)}]]$$

where $j > i$, $i = 1,2,\ldots,k$. Note that when this matrix is a matrix of 1s, we say that the two random variables are equal in mean form. When this matrix is $c \cdot \mathbf{1}$ for scalar $c > 0$ and **1** is a matrix of 1s, then we say that the two random variables are equal in mean shape.

Note that our model is invariant with respect to reflection. We feel that for specific biological problems this property can be of great advantage to the researcher. For example, invariance with respect to reflection makes geometrically based studies of asymmetry possible. In the study of archeological or paleontological samples, specimens are

often fragmentary. Assuming symmetry in the organisms under study, invariance with respect to reflection can allow direct, geometrically based comparison of fossils which have opposite sides preserved thereby increasing sample size.

TESTING FOR EQUALITY OF AVERAGE SHAPES

Suppose there are two populations whose shapes we want to compare. Let $\mathbf{X}_1$, $\mathbf{X}_2,\ldots,\mathbf{X}_n$ be a random sample of forms from Population I and $\mathbf{Y}_1, \mathbf{Y}_2,\ldots,\mathbf{Y}_m$ be a random sample from Population II. The null hypothesis is that the average shapes of the two populations are equal, which can be expressed using Definition 6, as follows:

$$H_0\text{: } \mathbf{F[E(X)]} = \mathbf{cF[E(Y)]} \text{ for some } c > 0$$

A natural way to test this hypothesis would be to estimate F[E(X)] and F[E(Y)] from the data, calculate the estimate of average form difference matrix D[E(X), E(Y)] using these estimates, and then test whether or not this matrix is "almost" a matrix of constants or not.

Estimating the form difference matrix

In the following, we offer two different procedures for estimating F[E(X)] and F[E(Y)].

Method 1

The most natural way to estimate F[E(X)] would be to estimate the average coordinates of X, E(X), and then calculate its form matrix. Although one can use any superimposition method, such as edge matching, we use generalized procrustes analysis (GPA) toward this end. Given $\mathbf{X}_1, \mathbf{X}_2,\ldots,\mathbf{X}_n$ we apply GPA (as described in Goodall and Bose, 1987) to get $\overline{\mathbf{X}}$. This $\overline{\mathbf{X}}$ is a consistent estimator of E(X) (see Goodall, 1991; however, see Lele, 1991b, for a further discussion). Similarly one can estimate E(Y) by $\overline{\mathbf{Y}}$. Here $\overline{\mathbf{X}}$ and $\overline{\mathbf{Y}}$ are coordinatewise averages of $\mathbf{X}_1$, $\mathbf{X}_2$, $\ldots,\mathbf{X}_n$ and $\mathbf{Y}_1, \mathbf{Y}_2,\ldots,\mathbf{Y}_m$ *after* superimposition. F[E(X)] and F[E(Y)] can then be estimated by using $\mathbf{F}(\overline{\mathbf{X}})$ and $\mathbf{F}(\overline{\mathbf{Y}})$.

Method 2

A natural and computationally simpler estimator of F[E(X)] is an average of the form matrices $\mathbf{F}(\mathbf{X}_1),\ldots,\mathbf{F}(\mathbf{X})_n)$, that is, an average of like linear distances within a sample of forms. Unfortunately, the resultant estimator is neither unbiased nor consistent for F[E(X)]. However the bias depends on the coefficient of variation and if it is small, the bias is negligible. We make this statement more precise in Theorem A1 in Appendix A. The theorem essentially shows that the form difference matrix calculated from the average $\mathbf{F}(\overline{\mathbf{X}})$ and $\mathbf{F}(\overline{\mathbf{Y}})$ is a consistent (or almost consistent) estimator of the true form difference matrix under fairly general conditions. For example, this result holds true even when the errors at various landmarks are dependent, nonidentical, or nonsymmetric around the origin.

Statistically, the behavior of the ratios can be tricky. We have studied the behavior of the estimator of the average form difference matrix described in Method 2 by simulation (Lele and Richtsmeier, 1991). The estimator is stable for moderate sample sizes.

In the following discussion we assume that F[E(X)] and F[E(Y)] are available using either of the above methods. For the sake of simplicity of notation we write these as $\mathbf{F}(\overline{\mathbf{X}})$ and $\mathbf{F}(\overline{\mathbf{Y}})$, respectively.

The next step is to calculate the average form difference matrix $\mathbf{D}(\overline{\mathbf{X}}, \overline{\mathbf{Y}})$ using definition 6. Thus

$$D_{ij}(\overline{\mathbf{X}}, \overline{\mathbf{Y}}) = F_{ij}(\overline{\mathbf{X}}) / F_{ij}(\overline{\mathbf{Y}}),\ i > j = 1,2,\ldots,K$$

In order to test the null hypothesis of similarity of form, we need to test whether or not this matrix is "too far" from a matrix of constants. Several test statistics can be proposed towards this objective: Let $\bar{D}$ be the average of the elements of the form difference matrix $D(\overline{X}, \overline{Y})$.

$$T_1 = \sum_{i,j} [D_{ij}(\overline{\mathbf{X}}, \overline{\mathbf{Y}}) - \bar{D}]^2 \tag{1}$$

$$T_2 = \sum_{i,j} |D_{ij}(\overline{\mathbf{X}}, \overline{\mathbf{Y}}) - \bar{D}| \tag{2}$$

$$T_3 = \max_{i,j} D_{i,j}(\overline{\mathbf{X}}, \overline{\mathbf{Y}}) - \min_{i,j} D_{ij}(\overline{\mathbf{X}}, \overline{\mathbf{Y}}) \tag{3}$$

$$T = \max_{i,j} D_{ij}(\overline{\mathbf{X}}, \overline{\mathbf{Y}}) / \min_{i,j} D_{ij}(\overline{\mathbf{X}}, \overline{\mathbf{Y}}) \tag{4}$$

If the null hypothesis is true, the first three quantities are close to zero, and the fourth quantity is close to one. We prefer the last statistic for the following mathematical and statistical reasons.

First, shape and shape difference are invariant under scaling. Let X and Y be two objects under consideration and let $c\mathbf{X}$ and $d\mathbf{Y}$ be their scaled versions with c not necessarily equal to d and $c > 0, d > 0$. We expect

that any test statistic that claims to indicate "shape difference" should assume identical values whether we compare X and Y or cX and dY. It is easy to check that T remains the same for both, but T_1, T_2, T_3 do not.

Second, calculation of the null distribution is uncomplicated because the null distribution is invariant under scaling. This follows from the above property.

Third, the test statistic is sensitive to changes in shape. There is a danger that the method may be too sensitive. To protect against spurious results one might robustify this statistic by applying some trimming. For example, one may take the ratio of the third quartile and the first quartile. However, as explained by Lele and Richtsmeier (1991), the extremes of the form difference matrix contain most of the information pertinent to form or shape difference. It seems imprudent to ignore the most useful information pertaining to the locus of shape difference in order to increase the robustness of the test. Routine statistical thinking in nonroutine problems can be very dangerous. We therefore do not suggest any robustified version of our proposed statistic, T. For small sample sizes, however, T can be somewhat unstable. We suggest that when faced with small sample sizes one should worry less about "acceptance" or "rejection" of the null hypothesis and consider the analysis exploratory. Studying the form difference matrix itself proves to be very useful in such situations (see Richtsmeier and Lele, 1990).

Fourth, this last test statistic results from the union–intersection principle. See Appendix B for details.

Estimation of the null distribution

Even in the simplest case of Gaussian perturbations, the analytical derivation of the null distribution of T is extremely complicated. Hence in the following we describe a bootstrap procedure for estimating the null distribution of the test statistic T. This is based on the well-known permutation test procedure coupled with Bootstrap (Efron, 1982) methodology to reduce the computational burden. A similar procedure for estimating the null distribution of a test statistic is employed by Clarren et al. (1987). See Romano (1988, 1989) for statistical justification of these procedures.

Bootstrap procedure

Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ and $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_m$ be the two samples. Let $Z = (\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}_{n+m})$, denote the mixed sample made up of X and Y.

Step 1. Select $\mathbf{Z}_i^*$, $i = 1,2,\ldots,n+m$ from Z randomly and with replacement.

Step 2. Split the bootstrap sample $\mathbf{Z}^* = (\mathbf{Z}_1^*, \mathbf{Z}_2^*, \ldots, \mathbf{Z}_{n+m}^*)$ in two groups $\mathbf{Z}_1^*, \ldots, \mathbf{Z}_n^*$ and $\mathbf{Z}_{n+1}^*, \ldots, \mathbf{Z}_{n+m}^*$ corresponding to the size of the original samples X and Y

Step 3. Calculate T^* for these two "samples", using the average form obtained by either Method 1 or 2. In our examples we have used Method 1.

Step 4. Repeat steps 1–3 B times where B is large (approximately 100). A histogram of T_j^*, $j = 1,2,\ldots,B$ estimates the null distribution of T, when H_0 is true.

Testing procedure

If the observed value of T, i.e., the value calculated with original samples X and Y is in the extreme right-hand tail of the null distribution, we reject H_0 at the appropriate level of significance. One may also report the P-value.

EXAMPLES

Craniofacial dysmorphology

Premature closure of craniofacial sutures (craniosynostosis) is a component of Crouzon and Apert syndromes. Irregularity of the pattern of premature craniosynostosis is common in both syndromes. In addition, these syndromes are marked by facial abnormalities, including shallow bony eye orbits, increased interorbital distance (hypertelorism), a "beaked," parrot-like nose, and defective formation of the maxilla resulting in a sunken appearance of the face. A more complete description of Crouzon and Apert craniofacial morphology can be found in Kreiborg (1986).

The data analyzed in the following example are coordinate locations of biological landmarks located on lateral X-rays of normal males and those affected with Apert and Crouzon syndrome. The 10 landmarks used in analysis are presented on an outline of a lateral projection of the skull as seen in an X-ray (Fig. 1) and are defined in Table 1. Details about the samples and data collection procedures can be found in Richtsmeier (1987).

Using Euclidean distance matrix analysis (EDMA) we compared a sample of four year old [$N_{(4)} = 20$] and 13-year-old [$N_{(13)} = 19$] normal males with age matched samples of Apert [$N_{(4)} = 5$; $N_{(13)} = 5$] and Crouzon [$N_{(4)} = 4$; $N_{(13)} = 5$] boys. In the comparison

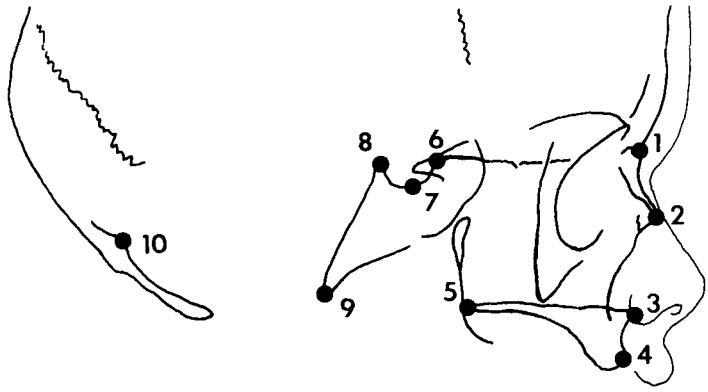

Fig. 1. Biological landmarks located on a two-dimensional representation of the human skull as seen in a lateral radiograph and used in analysis of normal, Crouzon, and Apert morphology.

TABLE 1. Definition and description of landmarks used as homologous points in analysis of normal, Crouzon, and Apert midsagittal craniofacial morphology

Landmark number	Landmark name and description
1	Nasion: Point of intersection of the nasal bones with the frontal bone
2	Nasale: Inferior-most point of intersection of the nasal bones
3	Anterior nasal spine: Anterior-most point at the medial intersection of the maxillary bones at the base of the nasal aperture
4	Intradentale superior: The point is located on the alveolar border of the maxilla between the central incisors
5	Posterior nasal spine: Posterior-most point of intersection of the maxillary bones on the hard palate
6	Tuberculum sella: "Saddle" of bone just posterior to the chiasmatic groove on the body of the sphenoid bone
7	Sella: Most inflexive point of the hypophyseal fossa. The hypophyseal (Pituitary) fossa is defined as the bony depression which holds the pituitary gland. This fossa is bounded by tuberculum sella anteriorly and posterior sella posteriorly
8	Posterior sella: A square plate of bone which serves as the posterior border of the hypophyseal fossa
9	Basion: The most anterior border of the foramen magnum
10	Internal occipital protuberance of the cruciate eminence of the occipital bone

of the 4-year-old normal males with the age-matched sample of Crouzon boys, the first step is to calculate a mean for each sample. To do this we applied a generalized procrustes algorithm to each sample separately. For each sample, linear distances between all possible pairs of points ($N = 45$) were computed from the suite of 10 averaged landmark locations. The resultant form matrix for 4-year-old normal boys and for 4-year-old boys with Crouzon syndrome were used to compute the form difference matrix. Like linear distances from the two form matrices were paired and a ratio was computed for each linear distance. In our example, linear distances from the normal sample serve as the numerator while linear distances from the Crouzon sample appear in the denominator. This matrix of ratios, the form difference matrix (Table 2), provides a distance by distance comparison of the average forms representing the two samples.

To test for difference between the two samples of forms, the statistic T is calculated ($T = 1.309/0.826 = 1.58$). The null distribution of T is calculated by first combining individuals from the normal sample and from the sample of boys with Crouzon syndrome ($N = 24$). From this combined sample, individuals are picked at random and

TABLE 2. Form difference matrices for the comparison of Apert and Crouzon with age-matched normal samples

Normal/Crouzon age 4		Normal/Apert age 4		Normal/Crouzon age 13		Normal/Apert age 13	
Ratio[1]	Landmarks[2]	Ratio	Landmarks	Ratio	Landmarks	Ratio	Landmarks
0.826	8-7	0.824	7-6	0.761	8-7	0.697	7-6
0.935	2-1	0.850	10-9	0.870	8-6	0.706	8-7
0.982	3-1	0.854	8-6	0.897	7-6	0.786	8-6
0.985	3-2	0.869	8-7	0.912	4-3	0.837	4-3
0.990	8-6	0.964	10-5	1.003	2-1	0.946	10-9
0.994	4-2	0.970	10-7	1.021	5-1	1.005	4-2
0.997	4-1	0.973	10-6	1.031	4-1	1.032	8-1
1.005	4-3	0.973	3-2	1.040	4-2	1.049	7-1
1.016	7-6	0.982	10-8	1.053	10-9	1.056	10-5
1.019	5-2	1.000	10-3	1.058	5-2	1.074	5-1
1.022	5-1	1.001	10-4	1.068	8-1	1.081	6-1
1.070	8-1	1.008	10-2	1.076	3-1	1.082	4-1
1.072	8-2	1.013	4-2	1.090	10-6	1.083	5-2
1.073	5-3	1.033	10-1	1.091	10-7	1.085	9-1
1.079	7-1	1.048	5-2	1.097	10-1	1.085	10-7
1.089	6-1	1.055	8-2	1.098	7-1	1.088	10-4
1.090	10-9	1.061	9-6	1.108	8-2	1.092	10-3
1.090	7-2	1.065	7-2	1.108	6-1	1.099	8-2
1.094	6-2	1.074	3-1	1.113	10-2	1.106	5-3
1.095	9-1	1.075	9-2	1.116	3-2	1.112	10-1
1.102	10-2	1.077	4-3	1.121	10-8	1.112	9-2
1.103	9-2	1.082	8-1	1.137	9-1	1.113	10-6
1.106	10-1	1.084	8-3	1.139	5-3	1.114	3-2
1.116	10-7	1.085	4-1	1.143	6-2	1.114	10-2
1.117	10-6	1.087	5-3	1.146	7-2	1.130	7-2
1.125	8-3	1.091	7-1	1.153	10-3	1.132	10-8
1.129	10-3	1.093	9-8	1.158	10-5	1.138	9-3
1.129	5-4	1.096	6-2	1.165	10-4	1.142	9-6
1.135	10-8	1.101	9-1	1.175	9-2	1.152	6-2
1.152	6-3	1.104	7-3	1.178	6-5	1.152	8.3
1.152	10-4	1.111	8-4	1.178	5-4	1.162	5-4
1.154	9-6	1.113	9-3	1.188	8-4	1.174	8-4
1.157	10-5	1.113	5-1	1.191	8-3	1.178	9-4
1.157	7-3	1.120	5-4	1.193	9-6	1.179	9-5
1.159	8-4	1.121	6-3	1.198	6-4	1.180	3-1
1.159	9-8	1.125	9-7	1.205	8-5	1.192	6-4
1.162	9-3	1.131	6-1	1.220	6-3	1.194	6-3
1.182	6-4	1.134	8-5	1.231	7-4	1.205	7-3
1.183	9-7	1.136	7-4	1.241	7-3	1.230	7-4
1.195	7-4	1.138	2-1	1.252	9-3	1.252	6-5
1.209	9-4	1.143	6-4	1.266	9-7	1.256	8-5
1.216	8-5	1.144	9-4	1.268	7-5	1.282	9-8
1.258	6-5	1.145	9-5	1.273	9-8	1.286	9-7
1.274	9-5	1.185	6-5	1.275	9-4	1.317	2-1
1.309	7-5	1.199	7-5	1.406	9-5	1.411	7-5

[1]Ratio i–j equals the distance between landmarks *i* and *j* in the normal group divided by the corresponding distance in the comparison group.
[2]Landmarks refer to the endpoints of each linear distance (see Table 1 for landmark numbers).

with replacement in order to form two samples of the size of the Crouzon and normal samples (i.e., $N = 20$ and $N = 4$). The comparison of these bootstrapped samples is done using the exact procedures outlined for comparing the original data. Mean forms are calculated, form matrices are computed and then compared by calculating a form difference matrix. T is calculated from the form difference matrix of the bootstrapped sample. This entire procedure is repeated 100 times and the resulting distribution of T is plotted as a histogram (Fig. 2). Since each individual form has an equal chance of being chosen during the bootstrap procedure, the composition of the bootstrapped samples is random. The location of T_{obs} with respect to the null distribution of T indicates the probability of obtaining T_{obs} when the sample forms are equal.

The P-value obtained in the comparison of normal with Crouzon at age 4 is 0.10. (See

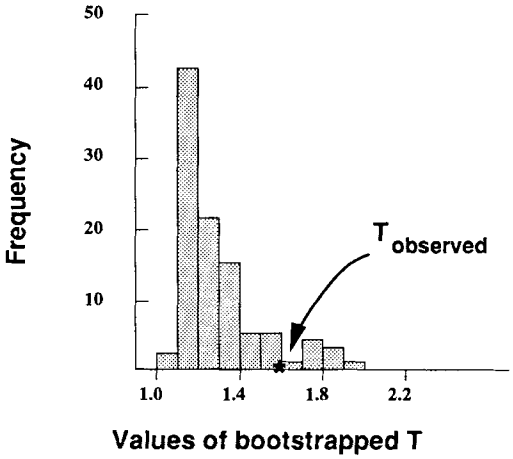

Fig. 2. Bootstrap estimate of the null distribution of T for the comparison of normal boys and those affected with Crouzon syndrome at age 4. T_{obs} was equal to 1.58 and 10% of the bootstrapped Ts exceeded T_{obs}.

Figure 2 for the bootstrap estimate of the null distribution of T.) Previous studies of Crouzon craniofacial morphology have noted a distinct dysmorphology local to the pituitary fossa (Kreiborg, 1976, 1986; Richtsmeier, 1987) and an extremely reduced posterior cranial base (Kreiborg and Pruzansky, 1981; Kreiborg, 1986). The posterior cranial base can be visualized on Figure 1 as that area between the pituitary fossa (landmarks 6, 7, 8) and basion (9). Our analysis supports previous observations, as landmarks 6, 7, and 8 are involved in many of the extreme ratios (see Table 2) and the distance from landmark 9 to landmarks 6, 7, and 8 are all at the maximum end of the ratio matrix. Following Bertelsen (1958), we feel that dysmorphology of the pituitary fossa is due to increased local bony resorption in response to intracranial pressure caused by craniosynostosis. The extreme dysmorphology local to the pituitary fossa results in a deepening of the fossa producing an *apparent* reduction in the length of the posterior cranial base.

Our analysis also indicates that the anteroposterior diameter of the occipital region of the neurocranium (measured as 10–9, 10–8) is smaller than normal in Crouzon syndrome. This is due to neurocranial dysmorphology associated with premature synostosis. In addition, palatal length, measured from landmark 4 to 5, is shorter in the Crouzon sample supporting previous observations of a smaller palate in Crouzon syndrome. The relationship of landmarks 4 and 5 with landmarks on the cranial base reflects the midfacial hypoplasia found in Crouzon syndrome and suggests the primacy of the posterior nasal spine in this regional dysmorphology.

Crouzon morphology ($N = 5$) is extremely different from normal ($N = 19$) at age 13 ($P = 0.01$). This suggests that the 13-year-old Crouzon morphology is more different from an age-matched normal sample than is the 4-year-old Crouzon form. These findings agree with those of Kreiborg and Pruzansky (1981) who found that the dysmorphology of Crouzon syndrome worsens with age (but see Richtsmeier, 1987 for dicussion). By age 13 the pituitary fossa (landmarks 6, 7, 8) is extremely enlarged in the Crouzon sample as indicated by ratios at the minimum end of the matrix. Reduction of distances measured from the cranial base to the face reflects the combination of an enlarged pituitary fossa and the persistent hypoplastic maxilla.

The P-value obtained in the comparison of normal ($N = 20$) with Apert syndrome ($N = 5$) at 4 years of age is 0.16. Like the younger Crouzon sample, the pituitary fossa is enlarged. We attribute this local dysmorphology to the same cause cited for the Crouzon sample: continual or increased intracranial pressure caused by craniosynostosis resulting in massive remodeling of the pituitary fossa. Like the Crouzon sample, distances measured from the palate (landmarks 3, 4, 5) to the cranial base (landmarks 6, 7, 8, 9) are reduced indicating maxillary hypolasia in 4-year-old Apert individuals.

By age 13, the Apert sample ($N = 5$) is distinct from the normal sample ($N = 19$), with a P-value of 0.01. Older Apert individuals are more different from age matched normals than are 4-year-old Apert individuals. Our results tend to support those which have characterized Apert syndrome as an age-progressive disease (Pruzansky, 1977; Kreiborg and Pruzansky, 1981; Richtsmeier, 1987). Age progressivity is specific to certain anatomical regions however, as a subgroup of the ratios is close to one in both age groups. Landmarks surrounding the pituitary fossa (6, 7, 8) are indicated as those which differ greatly between the normal and pathological samples at 13 years of age. The significance of this local pattern of dysmorphology associated with Crouzon and Apert syndromes should not be underestimated, nor should

TABLE 3. Landmarks used as homologous points in analysis of sexual dimorphism among adult Cebus apella

Landmark number	Landmark
8	Nasion
9	Nasale
10	Intradentale superior
11	Right premaxilla-maxilla junction at alveolus
12	Left premaxilla-maxilla junction at alveolus
13	Right frontal-zygomatic junction on orbital rim
14	Left frontal-zygomatic junction on orbital rim
15	Right zygomaxillare superior on orbital rim
16	Left zygomaxillare superior on orbital rim
17	Right zygomaxillare inferior
18	Left zygomaxillare inferior
31	Right maxillary tuberosity: maxillary-palatine intersection
32	Left maxillary tuberosity: maxillary-palatine intersection
35	Posterior nasal spine: vomer-palatine intersection
36	Vomer-sphenoid junction

the role of this pattern in the results of earlier studies based on registration systems centered on the pituitary fossa.

This analysis does not adequately describe the spectrum of craniofacial anomalies associated with Crouzon and Apert syndromes. This reflects a paucity of facial and neurocranial landmarks obtainable from a lateral X-ray, rather than flaws in our technique. To characterize the morphology of the Crouzon and Apert face, more data points from alternative (three-dimensional) sources are needed.

We stress that our results are to a large degree consistent with previous studies of craniofacial morphology of Apert and Crouzon syndrome. Additionally, our results underscore the extreme deformation of the pituitary fossa, an area that is frequently used as a registration center in the analysis of radiographs.

Sexual dimorphism in Cebus apella

The data analyzed in this example are three-dimensional coordinates of 15 biological landmarks (Table 3) located on the facial skeletons of male and female adult *C. apella* (Fig. 3). The numbering system for these landmarks is not continuous because these data are part of a larger study of craniofacial growth in New World monkeys (Corner and Richtsmeier, 1991a,b). The female sample ($N = 38$) was compared to the male sample ($N = 34$) using EDMA to determine the loci and magnitude of sexual dimorphism of *C. apella* faces. The P-value obtained is 0.01 indicating a significant degree of morphological distinction between adult male and female *C. apella* faces.

The form difference matrix for the comparison of female to male faces (Table 4) indicates that all linear distances are less than or nearly equal to 1. This demonstrates that the female face is generally smaller than the male *C. apella* face. The reader should note, however, that the form difference matrix is not a constant; sexual dimorphism of *C. apella* is not due to differences in size alone, but to differences in form. To clearly understand the nature of this dimorphism, close inspection of the form difference matrix is required.

Beginning our discussion at the maximum end of the matrix (Table 4), note that three of the linear distances that are nearly equal to 1 involve the maxillary tuberosities and posterior nasal spine, and measure the width of the posterior palate. The distances measured from maxillary tuberosities to the contralateral premaxillary–maxillary junction are also similar between the sexes indicating similar dimensions of the posterior palate in males and females along an oblique anteromedial axis.

The minimum end of the form difference matrix consists of those linear distances that are most different between the sexes. Distances measured from maxillary tuberosity to ipsilateral zygomaxillare inferior are most sexually dimorphic. Distances from the posterior nasale spine to zygomaxillare inferior and from left to right zygomaxillare inferior are also particularly dimorphic. These ratios represent sexual differences in the degree of flaring of the zygomatic regions among *C. apella* producing a wider male face.

Increased prognathism of the muzzle in males is evidenced by the ratios of distances measured from nasale to intradentale supe-

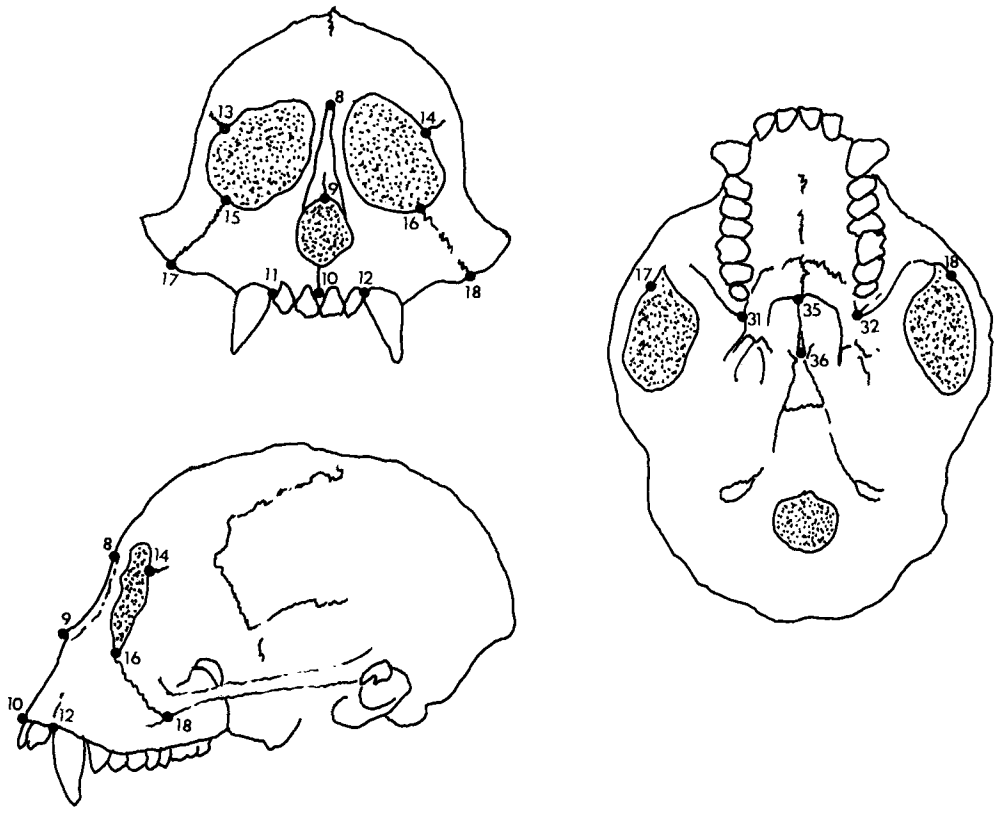

Fig. 3. Biological landmarks located on three planar views of the face of *C. apella*. Landmark locations were recorded in three dimensions using the 3Space digitizer.

rior (9, 10), from zygomaxillare superior to premaxillary–maxillary junction (16, 12 and 15, 11), from zygomaxillare superior to intradentale superior (16, 10 and 15, 10), and from premaxillary–maxillary junction to zygomaxillare inferior (18, 12 and 18, 11). This prognathism has both an anteroposterior and superoinferior component. Finally, distances measured from the vomer sphenoid junction (36) to points on the palate (posterior nasal spine, maxillary tuberosities, premaxillary–maxillary junction) indicate a fundamental sexual difference in the hafting of the face onto the basicranium among *C. apella*.

This example has demonstrated that although female *C. apella* facial skeletons are generally smaller than males, the difference is not a generalized isometric one. EDMA enables us to identify those loci that are most similar and most different between the sexes. Although the width of the posterior two-thirds of the palate is most similar between the sexes, the width of the midface, especially local to the zygomatic arches are the most sexually dimorphic structures. EDMA has also localized the sites of increased male prognathism of the muzzle and underscored a fundamental sexual difference in the hafting of the face onto the cranial base. Examination of linear distances with information that provides for geometric integrity of the forms considered (i.e. the form difference matrix) has allowed us to sort locations according to their contribution to sexual dimorphism of the facial skeleton.

TABLE 4. Form difference matrix for the comparison of female and male Cebus apella[1]

Female/male ratio	Landmarks
.8807	32-18
.8977	31-17
.9089	35-18
.9104	36-35
.9133	10-9
.9153	36-18
.9161	18-17
.9170	35-17
.9219	16-12
.9219	15-11
.9226	18-12
.9228	18-11
.9234	36-17
.9240	16-10
.9250	15-10
.9253	36-31
.9255	36-32
.9269	18-10
.9289	35-11
.9291	35-8
.9301	18-15
*	*
.9594	32-15
.9594	18-14
.9595	36-13
.9598	32-16
.9609	31-11
.9614	32-10
.9616	31-14
.9618	32-9
.9619	14-13
.9619	31-13
.9623	16-13
.9632	31-12
.9661	32-13
.9664	31-10
.9672	31-9
.9674	31-16
.9685	13-9
.9785	17-13
.9889	32-31
1.0128	35-32
1.0139	15-13
1.0141	35-31

[1]A total of 105 linear distances were computed; a subset of these, the extremal ends of the matrix, is shown.
*Indicates information missing.

DISCUSSION

In this paper we have shown how the Euclidean distance matrix-based approach for comparison of shapes suggested by Lele (1991a) can be extended for comparing average shapes of two groups statistically. We have illustrated its use in biological situations. Elsewhere we have compared the performance of this method with other methods theoretically (Lele, 1991a) and in a practical application (Richtsmeier and Lele, 1990).

A biologist is rarely interested in solely testing whether populations of forms or shapes are equal. Rather, the biologist seeks to identify those areas where the differences are most prominent. A ranking of various areas according to their contribution to the explanation of the shape differences observed is required to answer this question. In the companion paper (Lele and Richtsmeier, 1991) we suggest a method to address these issues.

ACKNOWLEDGMENTS

We thank Mr. Jingjang Liao and Mr. Warren Bilker for assistance in programming and Steve Danahey for help with data collection. Normative data come from the records of the Bolton-Brush Growth Study Center, Case Western Reserve University. The pathological samples come from the patient records of the Center for Craniofacial Anomalies, University of Illinois at Chicago (DE02872). We thank Dr. Richard Thorington for access to *C. apella* skulls from the collections of the National Museum of Natural History, Smithsonian Institution. This study was supported in part by grants from the Whitaker Foundation and the March of Dimes Birth Defects Foundation and the National Institutes of Health Biomedical Research Service Grants SO-7-RR05378 and 2 507 RR05445-27. Computer programs for these procedures are available from the authors when requests include a 5¼ in. formatted diskette. The authors are grateful to the Editor for his patience and kind encouragement. The referees' insightful comments are also acknowledged.

LITERATURE CITED

Bookstein FL (1986) Size and shape spaces for landmark data in two dimensions, Stat. Sci. *1*:181–242.

Bookstein FL (1989) Discussion of Kendall's paper. Stat. Sci. *4*:99–105.

Cheverud JM, and Richtsmeier JT (1986) Finite-element scaling applied to sexual dimorphism in rhesus macaque (*Macaca mulatta*) facial growth. Syst. Zool. *35*:381–399.

Clarren SK, Sampson PD, Larsen J, Donnell DJ, Barr HM, Bookstein FL, Martin DC, and Streissguth AP (1987) Facial effects of fetal alcohol exposure: Assessment by photographs and morphometric analysis. Am. J. Med. Genet. *26*:651–666.

Corner BD, and Richtsmeier JT (1991a) Morphometric analysis of craniofacial growth in *Cebus apella*. Am. J. Phys. Anthropol. *84*(3):323–342.

Corner BD, and Richtsmeier JT (1991b) Cranial growth in the squirrel monkey (*Saimiri sciureus*). Am. J. Phys. Anthropol. (in press).

Efron B (1982) The Jackknife, the Bootstrap and Other Resampling Plans. Philadelphia: SIAM.

Goodall C (1991) Procrustes methods in the statistical analysis of shapes. J. R. Stat. Soc. Ser. B 53, 285–339.

Goodall C, and Bose A (1987) Models and procrustes methods for the analysis of shape differences. Proceedings of the 19th Symposium on Interface Between Computer Science and Statistics, 86–92.

Kreiborg S (1986) Postnatal growth and development of the craniofacial complex in premature craniosynostosis. In, MM Cohen (ed.): Craniosynostosis: Diagnosis, Evaluation, and Management. New York: Raven Press.

Kreiborg S, and Pruzansky S (1981) Craniofacial growth in premature craniofacial synostosis. Scand. J. Plastic Reconstruct. Sur. *15*:171–186.

Lele SR (1991a) Some comments on coordinate free and scale invariant methods in morphometrics. Am. J. Phys. Anthropol. 85:407–418.

Lele SR (1991b) Discussion of Goodall's paper. J. R. Stat. Soc. Ser. B 53:334.

Lele SR, and Richtsmeier JT (1990) Statistical models in morphometrics: Are they realistic? Syst. Zool. *39*(1):60–69.

Lele SR, and Richtsmeier JT (1991) On comparing biological shapes: Detection of influential landmarks. Am. J. Phys. Anthropol. (in press).

Lewis JL, Lew WD, and Zimmerman JR (1980) A nonhomogeneous anthropometric scaling method based on finite element principles. J. Biomech. *13*:815–824.

Pruzansky S (1977) Time: The fourth dimension in syndrome analysis applied to craniofacial malformations. BD:OAS *13*:3–28.

Richtsmeier JT (1987) Comparative study of normal Crouzon, and Apert craniofacial morphology using finite element scaling analysis. Am. J. Phys. Anthropol. *74*:473–493.

Richtsmeier JT, and Lele SR (1990) Analysis of craniofacial growth in Crouzon syndrome using landmark data. J. Craniofac. Genet. Dev. Biol. *10*:39–62.

Romano J (1988) A bootstrap revival of some nonparametric distance tests. J. Am. Stat. Assoc. *83*:698–708.

Romano J (1989) Bootstrap and randomization tests of some nonparametric hypothesis. Ann. Stat. *17*:141–159.

Roth VL (1988) The biological basis of homology. In CJ Humphries, (ed.): Ontogeny and Systematics. New York: Columbia University Press.

Roy SN (1957) Some Aspects of Multivariate Analysis. New York: Wiley.

Van Valen LM (1982) Homology and causes. J. Morphol. *173*:305–312.

APPENDIX A

Let X and Y be two matrix valued random variables. Let $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ and $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_m$ be the random samples from X and Y, respectively.

Let $\mathbf{E}(\mathbf{X}_1), \mathbf{E}(\mathbf{X}_2), \ldots, \mathbf{E}(\mathbf{X}_n)$ and $\mathbf{E}(\mathbf{Y}_1), \ldots, \mathbf{E}(\mathbf{Y}_m)$ be the corresponding elementwise squared form matrices. Consider the squared distance between two landmarks i and j.

Let $E_{ij}(\mathbf{E}(\mathbf{X}))$ denote the squared distance between landmarks i and j in the average matrix of X. Similarly let $E_{ij}[\mathbf{E}(\mathbf{Y})]$ be the corresponding quantity for Y. Then $D_{ij}[\mathbf{E}(\mathbf{X})], [\mathbf{E}(\mathbf{Y})] = E_{ij}[\mathbf{E}(\mathbf{X})] / E_{ij}[\mathbf{E}(\mathbf{Y})]$.

By strong law of large numbers,

$$\frac{1}{n}\sum_{r=1}^{n} E_{ij}(\mathbf{X}_r) \to A_{ij} \qquad \text{a.s.}$$

$$\frac{1}{m}\sum_{r=1}^{m} E_{ij}(\mathbf{Y}_r) \to B_{ij} \qquad \text{a.s.}$$

as $n, m \to \infty$, where

$$A_{ij} = E[E_{ij}(\mathbf{X})] \neq E_{ij}[E(\mathbf{X})]$$

and

$$B_{ij} = E[E_{ij}(\mathbf{Y})] \neq E_{ij}[E(\mathbf{Y})]$$

It is trivial to check that

$$A_{ij} = E_{ij}[E(\mathbf{X})] + V_{ij}(\mathbf{X})$$

$$B_{ij} = E_{ij}[E(\mathbf{Y})] + V_{ij}(\mathbf{Y})$$

where $V_{ij}(\mathbf{X})$ and $V_{ij}(\mathbf{Y})$ are variances. An example makes this notation clear. Let $\mathbf{X}_1$ and $\mathbf{X}_2$ be two random variables with finite second moments. Then

$$\begin{aligned}&\mathbf{E}(\mathbf{X}_1 - \mathbf{X}_2)^{\mathrm{T}}(\mathbf{X}_1 - \mathbf{X}_2) = \\ &(\mu_1 - \mu_2)^{\mathrm{T}}(\mu_1 - \mu_2) \\ &+ \mathbf{E}\{(\mathbf{X}_1 - \mu_1) - (\mathbf{X}_2 - \mu_2)^{\mathrm{T}}[(\mathbf{X}_1 - \mu_1) \\ &- (\mathbf{X}_2 - \mu_2)]\}\end{aligned}$$

The first term corresponds to $E_{ij}(\mathbf{E}(\mathbf{X}))$ and the second term corresponds to $V_{ij}(\mathbf{X})$.

Theorem A1: If $V_{ij}(\mathbf{X}) / E_{ij}[\mathbf{E}(\mathbf{X})] = V_{ij}(\mathbf{Y}) / E_{ij}[\mathbf{E}(\mathbf{Y})]$, i.e., if the coefficients of variation are equal, then

$$\frac{A_{ij}}{B_{ij}} = \frac{E_{ij}[\mathbf{E}(\mathbf{X})]}{E_{ij}[\mathbf{E}(\mathbf{Y})]}$$

Proof:

$$\begin{aligned}&\frac{A_{ij}}{B_{ij}} - \frac{E_{ij}[\mathbf{E}(\mathbf{X})]}{E_{ij}[\mathbf{E}(\mathbf{Y})]} \\ &= \frac{V_{ij}(\mathbf{X})\,E_{ij}(\mathbf{E}(\mathbf{Y})) - V_{ij}(\mathbf{Y})E_{ij}(\mathbf{E}(\mathbf{X}))}{E_{ij}(\mathbf{E}(\mathbf{Y}))\,(E_{ij}(\mathbf{E}(\mathbf{Y})) + V_{ij}(\mathbf{Y}))} \\ &= \frac{\dfrac{V_{ij}(\mathbf{X})}{E_{ij}[\mathbf{E}(\mathbf{X})]} - \dfrac{V_{ij}(\mathbf{Y})}{E_{ij}[\mathbf{E}(\mathbf{Y})]}}{\dfrac{1}{E_{ij}[\mathbf{E}(\mathbf{X})]}\{E_{ij}[\mathbf{E}(\mathbf{Y})] + V_{ij}(\mathbf{Y})\}} \\ &= 0\end{aligned}$$

Hence the proof.

The condition of the theorem says that the larger the distance between two landmarks, the larger is the variation. This holds true for the model presented by Goodall and Bose (1987). We also note that Bookstein (1986) assumes that variation is very small as compared to the distances. Under this condition it is clear that the bias in our estimator is very small even when the condition of the theorem does not hold.

The theorem essentially says that the shape difference matrix can be estimated consistently using the average of the form matrices. Average of the form matrices is *not* a form matrix of the average object. However, for the purposes of testing and localising the shape differences it is sufficient to have a consistent estimator of the form difference matrix.

APPENDIX B
Union–intersection principle and derivation of the test

Roy (1957) introduced the union–intersection principle to develop tests particularly for multivariate distributions. The null hypothesis H_0, can frequently be expressed as an intersection of several null hypotheses H_{0a}, $a = 1,2,\ldots,h$. When expressed in this way, the null hypothesis, H_0, is supported if and only if *all* H_{0a}, $a = 1,2,\ldots,h$ are supported. For example, in the situation considered in this paper:

H_0: $D[\mathbf{E(X)}, \mathbf{E(Y)}] = c\ 1$ or equivalently
H_0: $D_{ij}\ [\mathbf{E(X)}, \mathbf{E(Y)}] = c$ for all
$i > j = 1,2,\ldots,k$

This hypothesis holds if and only if

$$H_{0;\,(i,j),\,(i',j')}\colon \frac{D_{ij}\,[\mathbf{E(X)}, \mathbf{E(Y)}]}{D_{i',j'}\,[\mathbf{E(X)}, \mathbf{E(Y)}]} = 1$$

is true for all pairs (i,j), (i',j') $i > j$, $i' > j'$. Thus

$$H_0 = \bigcap_{\substack{(i,j)\\(i',j')}} H_{0;\,(i,j),\,(i',j')}$$

This is the intersection part of the union–intersection principle. In words, this says that two shapes are equal if and only if *any* two elements in the form difference matrix are identical.

The union part of the union–intersection principle requires that we accept H_0 if and only if *all* the subhypotheses $H_{0;\,(i,j),\,(i',j')}$ are accepted, or equivalently, reject H_0 if *any one* of the subhypotheses is rejected.

Note that the most different pair of elements in the form difference matrix gives the ratio

$$\frac{\max\limits_{i,j} D_{ij}\,[\mathbf{E(X)}, \mathbf{E(Y)}]}{\min\limits_{i,j} D_{ij}\,[\mathbf{E(X)}, \mathbf{E(Y)}]}$$

which is consistently estimated by

$$T = \max_{i,j} D_{ij}\,(\overline{\mathbf{X}}, \overline{\mathbf{Y}}) / \min_{i,j} D_{ij}\,(\overline{\mathbf{X}}, \overline{\mathbf{Y}})$$

If this ratio is not "too far" from 1, all the other subhypotheses corresponding to other pairs cannot be "too far" from 1. Hence we accept H_0 if T is "close" to 1 otherwise we reject H_0. The proposed test is thus a union–intersection test.