Communications for Statistical Applications and Methods
2023, Vol. 30, No. 1, 37–48
https://doi.org/10.29220/CSAM.2023.30.1.037
Print ISSN 2287-7843 / Online ISSN 2383-4757

Estimation of missing landmarks in statistical shape analysis

Sang Min Shin[a], Jun Hong Kim[a], Yong-Seok Choi[1,a]

[a]Department of Statistics, Pusan National University, Korea

Abstract

Shape analysis is a method for measuring, describing and comparing the shape of objects in geometric space. An important aspect is to obtain Procrustes distance based on least square method. We note that the shape is all the geometrical information that remains when location, scale and rotational effects are filtered out from an object. However, and unfortunately, when we cannot measure some landmarks which are some biologically or geometrically meaningful points of any object, it is not possible to measure the variation of all shapes of an object, including that of the incomplete object. Hence, we need to replace the missing landmarks. In particular, Albers and Gower (2010) studied the missing rows of configurations in Procrustes analysis. They noted that the convergence of their approach can be quite slow. In this study, alternatively, we derive an algorithm for estimating the missing landmarks based on the pre-shapes. The pre-shape is invariant under the location and scaling of the original configuration with the centroid size of the pre-shape being one. Therefore we expect that we can reduce the amount of total computing time for obtaining the estimate of the missing landmarks.

Keywords: estimate, missing landmark, pre-shape, Procrustes distance, shape analysis

1. Introduction

Statistical shape analysis is an area of study arising from a wide variety of applications, including the fields of archaeology, biology, chemistry, geography, image analysis and medicine among others. It is only within the last decade that satisfactory techniques have emerged (Dryden and Mardia, 1993). In statistical shape analysis, an object can be specified by a geometrical shape obtained from a set of points referred to as landmarks. Dryden and Mardia (1998) defined that a landmark is a point of correspondence on each object that matches between and within populations. They also define a configuration as a set of landmarks on a particular object. The phrase "shape of an object" is defined as all the geometrical information that remains when location, scale and rotational effects are filtered out from the configuration (Kendall, 1984). That means, the shape is invariant under the transformations of translation, rotation and scaling.

Typical aims of statistical shape analysis include estimating a mean shape, investigating variability in shapes, testing differences in mean shapes between two or more groups and describing these differences. But sometimes configurations can have one or more missing landmarks because of the partial damage of objects or the loss of data. Therefore we need to replace the missing landmarks. Moreover, in Procrustes analysis which is the foundation of shape analysis, missing rows of configuration matrices were studied by Commandeur (1991). And Ten Berge *et al.* (1993) extended this to

[1] Corresponding author: Department of Statistics, Pusan National University, 2 Busandaehak-ro, Geumjeong-Gu, Busan 46241, Korea. E-mail: yschoi@pusan.ac.kr

Published 31 January 2023 / journal homepage: http://csam.or.kr

an unpatterned allocation of missing values, but both for the important special case where the transformations are orthogonal. Gower and Dijksterhuis (2004) addressed the problem more generally and Albers and Gower (2010) introduced an approach useful for handling the missing values of configuration matrices in the context of general linear transformations, but they had the problem of slow convergence in the process of estimating missing landmarks. In this study, we will adapt and modify their process of shape analysis for estimating shapes of incomplete configurations.

2. Preliminary

In shape analysis, the Procrustes analysis involves matching configurations with transformations of translating, scaling and rotating to be as close as possible to Euclidean distance using least squares techniques. The configuration matrix is defined as a $k \times m$ matrix of Cartesian coordinates of k-landmarks in m-dimensions for general m with $k > m$. Thus, we can consider the case where two configuration matrices, X_1 and X_2 are available. For fitting and comparing them, ordinary Procrustes analysis (OPA) involves the least squares matching of X_1 onto X_2. Examination of the matching parameters α, β and Γ is carried out by minimizing the squared Euclidean distance

$$D^2{}_{\text{OPA}}\,(X_1, X_2) = \left\| \left(\beta X_1 \Gamma + 1_k \alpha^t\right) - X_2 \right\|^2, \tag{2.1}$$

where $\|X\| = \sqrt{tr\,[X^t X]}$ is the Euclidean norm, α is an $m \times 1$ vector, $\beta > 0$ is a scale parameter and Γ is an $m \times m$ orthogonal rotation matrix. Without loss of generality we assume that the configuration matrices X_1 and X_2 have been centered. Then the OPA solution to the minimization of Equation (2.1) is given by

$$\hat{\alpha} = 0, \quad \hat{\Gamma} = VU^t, \quad \hat{\beta} = \frac{tr\left[X_2^t X_1 \hat{\Gamma}\right]}{tr\left[X_1^t X_1\right]},$$

where $X_2^t X_1 = \|X_1\|\,\|X_2\|\,U \Lambda V^t$ and both U and V are orthogonal matrices satisfying $U^t U = V^t V = I_m$ and $\Lambda = \text{diag}\,(\lambda_1, \ldots, \lambda_m)$ is a diagonal matrix of singular values $\lambda_1 \geq \cdots \geq \lambda_m \geq 0$. Now consider the general case where $n \geq 2$ configuration matrices, $X_1, \ldots, X_n$ are available. The generalized Procrustes analysis (GPA) involves translating, scaling and rotating the configurations relative to each other so as to minimize the total sum of squares. We minimize a quantity proportional to the sum of squared norms of pairwise differences,

$$D^2{}_{\text{GPA}} = \frac{1}{n} \sum_{i<j}^{n} \left\| \left(\beta_i X_i \Gamma_i + 1_k \alpha_i^t\right) - \left(\beta_j X_j \Gamma_j + 1_k \alpha_j^t\right) \right\|^2, \tag{2.2}$$

where α_i $(i = 1, \ldots, n)$ is an $m \times 1$ location parameters, $\beta_i > 0$ is a scale parameter and Γ_i is an $m \times m$ orthogonal rotation matrix. Clearly, Equation (2.2) is minimized by choosing all β_i to be close to zero. Apart from this trivial solution, Gower(1975) notices a constraint such that $\sum_{i=1}^n \|\beta_i X_i\|^2 = \sum_{i=1}^n \|X_i\|^2$. Here, Equation (2.2) may be written alternatively as

$$D^2{}_{\text{GPA}} = \left\| \left(\beta_i X_i \Gamma_i + 1_k \alpha_i^t\right) - \mu_x \right\|^2, \tag{2.3}$$

where $\mu_x = 1/n \sum_{i=1}^n (\beta_i X_i \Gamma_i + 1_k \alpha_i^t)$ is the average configuration. Thus the optimal solution to the minimization of Equation (2.3) over μ_x is given by $\hat{\mu_x} = 1/n \sum_{i=1}^n (\hat{\beta}_i X_i \hat{\Gamma}_i + 1_k \hat{\alpha}_i^t)$. To obtain $\hat{\alpha}_i$, $\hat{\beta}_i$ and

$\hat{\Gamma}_i$, Gower (1975), Ten Berge (1977) and Kent (1994) have given different definitions and algorithms. However there is an equivalence among these definitions that yield the same Procrustes mean shape.

Albers and Gower (2010) proposed an algorithm for handling missing values of a configuration matrix. This is related with the algorithm for estimating the missing landmarks of the configuration with the transformation of centering, scaling and orthogonal rotating. Furthermore, since Procrustes methods are based on the least squares technique, the algorithm uses least squares method. Here, we need to describe the algorithm proposed by Albers and Gower (2010).

Recall the GPA criterion of Equation (2.3) and let $\mu_x^{(h)}$ be the average of configurations after the configurations have been translated, rotated and scaled excluding the h^{th} configuration, such that

$$\mu_x^{(h)} = \frac{1}{n-1} \sum_{i \neq h}^{n} \left(\beta_i X_i \Gamma_i + 1_k \alpha_i^t \right).$$

Then the Procrustes sum of squares of Equation (2.3) is rewritten as

$$D^2{}_{\text{GPA}} = \sum_{i \neq h}^{n} \left\| \left(\beta_i X_i \Gamma_i + 1_k \alpha_i^t \right) - \mu_x^{(h)} \right\|^2 + \frac{n-1}{n} \left\| \left(\beta_h X_h \Gamma_h + 1_k \alpha_h^t \right) - \mu_x^{(h)} \right\|^2. \tag{2.4}$$

Now, suppose X_k has $r\,(< k)$ missing landmarks. If the rm missing cells of X_h contain putative values, then we seek to update them by minimizing the criterion of Equation (2.4). Denote Δ by an updating matrix with zeros everywhere except for the cells corresponding to the missing values in X_h and let $X_h - \Delta$ be the estimated configuration matrix of X_h. If we want to update the h^{th} configuration from X_h to $X_h - \Delta$, it suffices to consider minimizing over Δ the term

$$\left\| \beta_h \left(I_k - K \right) \left(X_h - \Delta \right) \Gamma_h - \mu_x^{(h)} \right\|^2, \tag{2.5}$$

where I_k is the $k \times k$ identity matrix, $K = (1/k) 1_k 1_k^t$. Then we can rewrite Equation (2.6) as

$$\left\| \beta_h \left(I_k - K \right) \Delta \Gamma_h - \left(\beta_h \widetilde{X}_h \Gamma_h - \mu_x^{(h)} \right) \right\|^2, \tag{2.6}$$

where $\widetilde{X}_h = (I_k - K) X_h$ which is called the centered shape of X_h. Equation (2.6) is itself a Procrustes problem where now it is Δ, rather than β_h and Γ_h. And let $Y = \beta_h \widetilde{X}_h - \mu_x^{(h)} \Gamma_h^t$ to minimize Equation (2.6), which is equal to minimize,

$$\beta_h^2 tr \left[\Delta^t \left(I_k - K \right) \Delta \right] - 2 \beta_h tr \left[\Delta Y^t \right]. \tag{2.7}$$

Equation (2.7) is to be minimized over Δ, which contains only rm active values, $\delta = (\delta_1, \ldots, \delta_{rm})^t$. So, for $rm \times rm$ matrix T and rm-vector y, Equation (2.7) is a quadratic form in δ which may be written as

$$\beta_h^2 \delta^t \left(I_{rm} - T \right) \delta - 2 \beta_h \delta^t y + \text{constant}, \tag{2.8}$$

where y consists of the rm-values of Y corresponding to the missing value position in Δ and δ. Thus, Equation (2.8) is minimal when

$$\delta = \frac{1}{\beta_h} \left(I_{rm} - T \right)^{-1} y.$$

Normally, Δ will have some zero rows and it is only necessary to examine the non-zero part. Therefore, by ordering δ by columns of Δ it will be found that T becomes block-diagonal with diagonal blocks of the form $(1/k)\ 1_r\ 1_r^t$, which correspond to a column with r missing values.

This updating process in X_h will estimate the missing values such that $\beta_h \widetilde{X}_h \Gamma_h$ is closest to $\mu_x^{(h)}$. After this change in X_h, a new generalized Procrustes step is needed to find the optimal transformations of X_i, $i = 1, \ldots, n$ including X_h and to re-estimate $\mu_x^{(h)}$. If we repeat the updating process and the Procrustes step until the Procrustes sum of squares, Equation (2.3) cannot be reduced further, and we finally obtain the estimate of the missing values of X_h.

3. Estimation of missing landmarks in shape analysis

We already mentioned that a shape is all the geometrical information that remains when location, scale and rotational effects are filtered out from the configuration. To remove the location and scale effects from a configuration matrix X, we can use a pre-shape in shape analysis. Note that the term pre-shape was first used by Kendall (1984). The pre-shape of a configuration matrix X is given by

$$Z = \frac{\mathrm{HX}}{\|\mathrm{HX}\|},$$

where H, which is called the Helmert sub-matrix, is a $(k-1) \times k$ matrix with orthonormal rows that satisfies $HH^t = I_{k-1}$ and $H^t H = I_k - K$. In particular, the l^{th} row of the Helmert sub-matrix H is given by

$$(h_l, \ldots, h_l, -lh_l, 0, \ldots, 0), \quad h_l = -\{l(l+1)\}^{-\frac{1}{2}}; l = 1, \ldots, k-1,$$

where the l^{th} row consists of h_l repeated l-times followed by $-lhl$ and then k–l–1 zeros. The transformed HX does not depend on the original location of X and $\|\mathrm{HX}\| = \|\widetilde{X}\|$ which is the centroid size of X, so that the pre-shape is invariant under the location and scaling of X. And the pre-shape space is a hypersphere of unit radius in $(k-1)m$ real dimensions since $\|Z\| = 1$. Also, in order to remove the rotational effect from configuration X, we identify all rotated versions of the pre-shape with each other and this set is the shape of X.

When a configuration is incomplete, we cannot obtain the shape information of the configuration. Therefore, we need to estimate the missing landmarks of the incomplete configuration. We will adapt and modify the algorithm of Albers and Gower (2010) to estimate the missing landmarks in statistical shape analysis. At first, as with the previous section we assume that the cells with unknown values of X_h contain some putative values that we seek to update by minimizing the GPA criterion. Since the Euclidean geometry of the configuration space induces spherical geometry in pre-shape space, the squared Procrustes distance between X_h and $\mu_x^{(h)}$ is defined as

$$d_P^2\left(X_h, \mu_x^{(h)}\right) = \inf_{\Gamma_h} \left\|Z_h \Gamma_h - \mu_z^{(h)}\right\|^2, \tag{3.1}$$

where $\mu_z^{(h)} = H\mu_x^{(h)}/\|H\mu_x^{(h)}\|$ and $Z_h = \mathrm{HX}_h/\|\mathrm{HX}_h\|$ are the pre-shapes of $\mu_x^{(h)}$ and X_h, respectively. Note that, the estimate of the minimizing rotation parameter Γ_h can be obtained by the OPA matching Z_h onto $\mu_z^{(h)}$.

Now, we want to update the incomplete configuration from X_h to $X_h - \Delta$ using the squared Procrustes distance. Before we estimate the rm missing values of X_h, we should denote the centroid size of the estimates of X_h as $1/c = \|H(X_h - \Delta)\|$. In addition, since the pre-shape space is a hypersphere of unit radius, we have a constraint in the following

$$\|cH(X_h - \Delta)\|^2 = 1. \tag{3.2}$$

Here, since $H^t H = I_k - K$ is idempotent, Equation (3.2) can be rewritten as

$$c^2 \left\|\widetilde{X_h}\right\| - 2c\zeta^t x + \zeta^t (I_{rm} - T)\zeta = 1, \tag{3.3}$$

where $\zeta = c\delta$ and x is an $rm \times 1$ vector which consists of the rm-values of $\widetilde{X_h}$ corresponding to the missing value position in Δ and δ.

Now, to estimate the r missing landmarks of X_h, we can consider minimizing the squared Procrustes distance in Equation (3.1) over Δ as follows

$$\left\|cH(X_h - \Delta)\Gamma_h - \mu_z^{(h)}\right\|^2, \tag{3.4}$$

and Equation (3.4) is rewritten as

$$c^2 \left\|\widetilde{X_h}\right\|^2 - 2c\zeta^t x + \zeta^t J\zeta + 2\zeta^t w - 2ctr\left[X_h^t \widetilde{\mu_z}^{(h)} \Gamma_h^t\right] + \left\|\mu_z^{(h)}\right\|^2, \tag{3.5}$$

where $J = I_{rm} - T$ and $\widetilde{\mu}_z^{(h)} = H^t \mu_z^{(h)}$ is the centered pre-shape of $\mu_x^{(h)}$. And w consists of the rm-values of $\widetilde{\mu}_z^{(h)} \Gamma_h^t$ corresponding to the missing value position in Δ and δ. Next, defining with matrix notations

$$\theta = \begin{pmatrix} \zeta \\ c \end{pmatrix}, \quad A = \begin{pmatrix} J & -x \\ -x^t & \left\|\widetilde{X}_h\right\|^2 \end{pmatrix}, \quad a = \begin{pmatrix} -w \\ tr\left[X_h^t \widetilde{\mu}_z^{(h)} \Gamma^t\right] \end{pmatrix}$$

we finally complete the square form of Equation (3.5) such that

$$\theta^t A\theta - 2\theta^t a + \text{constant} \tag{3.6}$$

and having a constraint $\theta^t A\theta = 1$ from Equation (3.3). Since A is positive definite, the optimal solution of Equation (3.6) is given by

$$\hat{\theta} = \begin{pmatrix} \eta J^{-1}x - J^{-1}w \\ \eta \end{pmatrix} \sqrt{\frac{1}{\eta^2/\nu + w^t J^{-1} w}}\,, \tag{3.7}$$

where $\nu = (\|\widetilde{X}_h\|^2 - x^t J^{-1} x)^{-1}$ and $\eta = \nu tr[X_h^t \widetilde{\mu}_z^{(h)} \Gamma^t] - \nu x^t J^{-1} w$.

As with the algorithm of Albers and Gower (2010), we additionally need a general Procrustes step to find the optimal transformations of X_i, $i = 1, \ldots, n$ including X_h after the change in X_h. However, the Procrustes sum of squares in their algorithm converges very slowly due to the scale parameter β_i. Because of the change of the centroid size of X_h in each loop, all of β_i may be changed in each Procrustes step. As a result, there is variation of the Procrustes sum of squares in each loop. Therefore to investigate variability in all of the shapes, we use the sum of squared Procrustes distances such that

$$D_P^2 = \sum_{i=1}^{n} \|Z_i \Gamma_i - \mu_z\|^2\,, \tag{3.8}$$

where $\mu_z = H\mu_x / \|H\mu_x\|$ is the pre-shape of μ_x. If we repeat the updating process and the Procrustes step until Equation (3.8) is converged, we finally obtain the estimate of the missing landmarks of X_h.

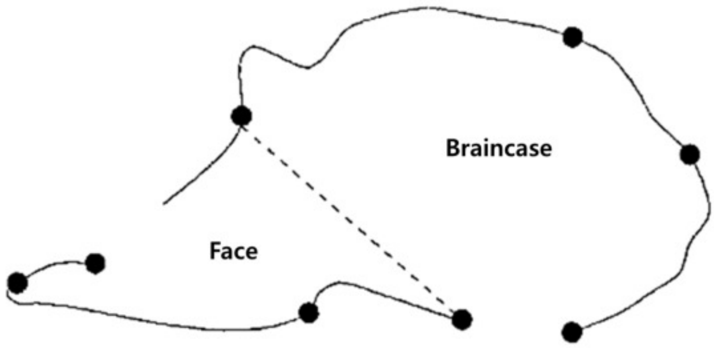


Figure 1: *Eight landmarks on a gorilla skull (Dryden and Mardia, 1998).*

Our proposed updating process needs to compute two parameters, ζ and c for obtaining δ. Therefore, in each updating process, the required computing time is more than that of Albers and Gower (2010). However, we use pre-shapes instead of configurations when we investigate variation of shapes. The pre-shape is invariant under the location and scale of the original configuration and the centroid size of the pre-shape is one. Thus, Equation (3.8) is less than that of Equation (2.3) in the variation of the squared Procrustes distances because we only consider the rotation of the pre-shape. Hence, we can reduce the amount of total computing time in finally obtaining the estimate of the missing landmarks of X_h. Here, we summarize our **Algorithm 1** as follows.

Algorithm 1: algorithm

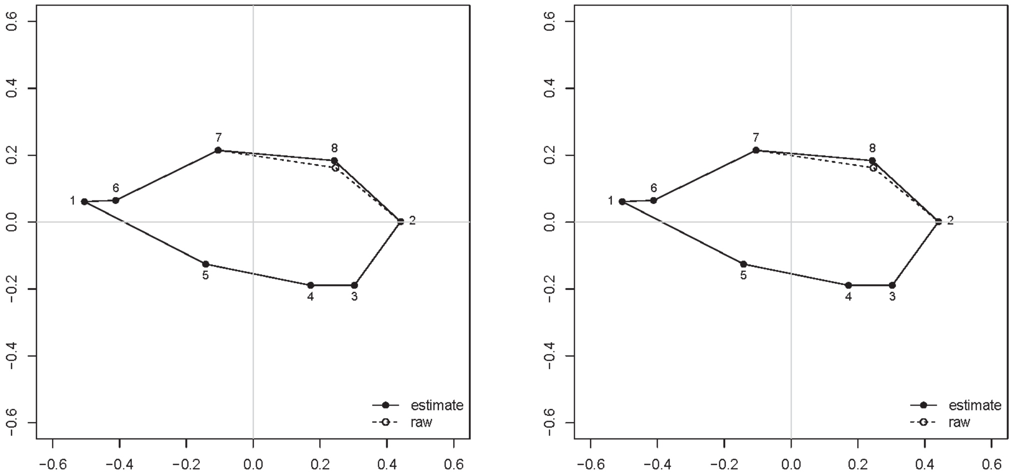
Step 1 : (Initial Procrustes step) For given configurations, compute $\mu_x^{(h)}$ using the GPA algorithm. And set missing landmarks in X_h to some initial values.
Step 2 : (Updating Process) For current pre-shape Z_h and $\mu_z^{(h)}$, estimate Γ_h by OPA matching Z_h onto $\mu_z^{(h)}$, and use $\hat{\Gamma}_h$ to compute $\theta = (\zeta^t, c)^t$ in Equation (3.7). Then, calculate the updating values $\delta = \zeta/c$ for the missing values of X_h and renew X_h. Repeat this step until Equation (3.5) cannot be reduced further.
Step 3 : (Procrustes step) Recompute the Procrustes fits $X_i^p = \beta_i \widetilde{X}_i \Gamma_i$ $(i = 1, \ldots, n)$ according to the GPA algorithm, including X_h. Then, the new $\mu_x^{(h)}$ is obtained.
Step 4 : (Repetition) Repeat steps 2 and 3 until the squared Procrustes distances of Equation (3.8) cannot be reduced further.

4. Examples

4.1. Estimation for the gorilla skulls

Consider the gorilla skulls data described in detail by O'Higgins (1989) and O'Higgins and Dryden (1993). There are 29 male and 30 female adult gorilla skulls with eight landmarks shown in Figure 1. In this paper, without loss of generality we assume that the configurations have been centered and scaled.

In the first example, we make each landmark of a configuration as a missing landmark and then re-estimate it. We performed a total of $8 \times 59 = 472$ re-estimates. In the process of re-estimating, we assume that the class of a given incomplete configuration is known. Thus, the GPA is available to find



(a) (b)

Figure 2: *Estimated configurations by (a) Algorithm I and (b) Algorithm II.*

the estimate of the mean shape for each class excluding the incomplete configuration. In this way, we try to compare the performance of our proposed algorithm with that of Albers and Gower (2010). The evaluation of an estimate of each configuration is measured as the square root of the residual sum of squares (SQRSS) which is Euclidean norm of the difference between raw configuration X_{raw} and the estimate of incomplete configuration $\hat{X}_{\text{mis}}$ such that

$$\mathbf{SQRSS} = \left\| X_{\text{raw}} - \hat{X}_{\text{mis}} \right\|. \tag{4.1}$$

For the convenience of notation, let us denote the algorithms as follows :

- Algorithm I : The algorithm of Albers and Gower (2010).
- Algorithm II : The proposed algorithm given by **Algorithm 1**.

Figure 2 shows the estimated configuration of the landmark 8 of a male gorilla skull. It is not easy to identify the difference between estimated configurations. Actually, the estimates of the two algorithms are almost completely matched and the SQRSS is calculated as 0.0219 in both cases. Table 1 gives the averages of the SQRSS (MSQRSS) for the cases of estimating each landmark for female and male gorilla skulls. We can also confirm that the MSQRSS of the two algorithms are equivalent, but convergence time is different. The ratio of the elapsed time for estimating the missing landmark using two algorithms is calculated as 36.26. In other words, our algorithm could estimate the missing landmark about 36.3 times faster than that of Albers and Gower (2010). Table 2 gives the averages of the ratio of the elapsed time for estimating each landmark. We can also confirm that the elapsed time of Algorithm II is extremely shorter than that of Algorithm I.

In the second example, for the randomly selected two of eight landmarks, we make them as missing landmarks and then we re-estimate them. Then, we performed a total of ${}_8C_2 \times 59 = 1{,}652$ re-estimates. Table 3 gives the MSQRSS for the cases of estimating each pair of landmarks for female and male gorilla skulls. In the table, the MSQRSS is presented for only one case because the estimates of Algorithm I and II are almost completely equivalent.

Table 1: The MSQRSS of Algorithm I and II for gorilla skulls with one deleted landmark

Deleted landmark	Female		Male	
	Algorithm I	Algorithm II	Algorithm I	Algorithm II
1	0.0219	0.0219	0.0245	0.0245
2	0.0233	0.0233	0.0280	0.0280
3	0.0190	0.0190	0.0197	0.0197
4	0.0189	0.0189	0.0173	0.0173
5	0.0146	0.0146	0.0170	0.0170
6	0.0126	0.0126	0.0163	0.0163
7	0.0162	0.0162	0.0210	0.0210
8	0.0226	0.0226	0.0253	0.0253

Table 2: The average ratio of the elapsed time of Algorithm I and II for gorilla skulls with one deleted landmark

	Deleted landmark							
	1	2	3	4	5	6	7	8
Female	46.46	45.04	44.71	45.97	40.58	44.91	41.48	46.50
Male	44.74	46.40	44.35	44.01	42.15	44.84	41.81	44.16

Table 3: The MSQRSS of Algorithm I and II for gorilla skulls with two deleted landmarks

Deleted landmarks	Female	Male	Deleted landmarks	Female	Male
1, 2	0.0318	0.0378	3, 5	0.0252	0.0286
1, 3	0.0295	0.0315	3, 6	0.0235	0.0263
1, 4	0.0287	0.0292	3, 7	0.0265	0.0306
1, 5	0.0272	0.0290	3, 8	0.0283	0.0334
1, 6	0.0406	0.0517	4, 5	0.0261	0.0270
1, 7	0.0283	0.0323	4, 6	0.0235	0.0242
1, 8	0.0321	0.0364	4, 7	0.0266	0.0284
2, 3	0.0337	0.0396	4, 8	0.0304	0.0317
2, 4	0.0288	0.0342	5, 6	0.0208	0.0241
2, 5	0.0282	0.0338	5, 7	0.0240	0.0307
2, 6	0.0274	0.0329	5, 8	0.0285	0.0325
2, 7	0.0305	0.0371	6, 7	0.0206	0.0257
2, 8	0.0404	0.0422	6, 8	0.0264	0.0308
3, 4	0.0352	0.0329	7, 8	0.0310	0.0379

The overall values of the MSQRSS increased more than the values in Table 1. In particular, when a configuration is missing on landmark 1 and 6, the MSQRSS is maximum for both female and male. Figure 3 shows an example of these cases. The configuration in Figure 3 is a estimated male gorilla skull and the SQRSS of the configuration is calculated as 0.0581. The estimates of these landmarks have a large bias because two missing landmarks are adjacent and especially landmark 1 is the extreme curvature on the edge of the configuration.

Table 4 shows the average ratio of the elapsed time for estimating each pair of landmarks. From this table, the elapsed time of Algorithm II is more than 45 times faster than Algorithm I. It can be confirmed again that the elapsed time of our algorithm is extremely shorter than that of Albers and Gower (2010).

4.2. Estimation for the three-dimensional configurations of macaque skulls

Consider the macaque skulls data described in Dryden and Mardia (1993) obtained the random samples of 9 male and 9 female skulls. There exists a total of 26 landmarks for each skull, and 7 landmarks

Table 4: The average ratio of the elapsed time of Algorithm I and II for gorilla skulls with two deleted landmarks

Deleted landmarks	Female	Male	Deleted landmarks	Female	Male
1, 2	48.94	45.87	3, 5	45.42	45.82
1, 3	48.38	43.54	3, 6	46.87	46.89
1, 4	45.78	46.05	3, 7	44.30	42.57
1, 5	46.87	44.84	3, 8	44.40	43.16
1, 6	47.82	47.22	4, 5	46.08	45.68
1, 7	47.27	43.85	4, 6	46.78	44.55
1, 8	45.98	44.52	4, 7	43.41	43.53
2, 3	47.84	46.65	4, 8	43.41	42.74
2, 4	44.88	46.31	5, 6	47.52	43.78
2, 5	43.33	46.30	5, 7	42.51	42.62
2, 6	46.41	47.06	5, 8	44.77	40.58
2, 7	46.27	45.11	6, 7	44.34	46.47
2, 8	46.87	46.83	6, 8	45.14	46.61
3, 4	48.56	47.47	7, 8	45.93	46.83

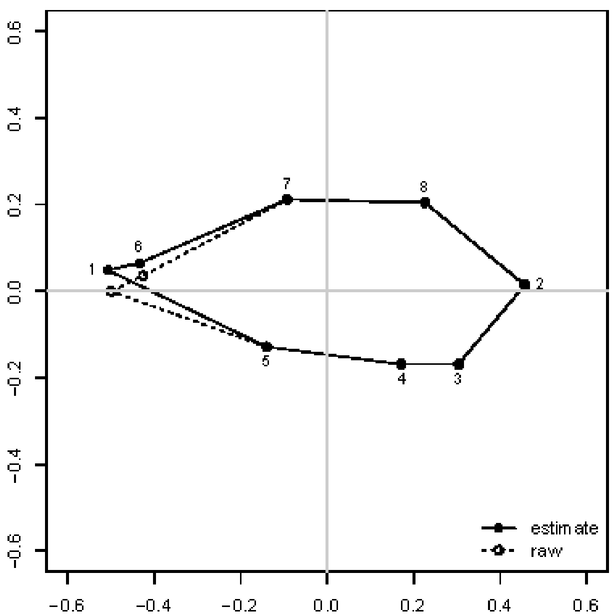


Figure 3: *Estimate of landmarks 1 and 6 of a male gorilla skull.*

in Figure 4 are taken 67 for analysis. The seven chosen landmarks are prosthion (1), opisthion (2), bregma (3), nasion (4), asterion (5), midpoint of temp suture (6) and interfrontomalare (7).

For each skull, we randomly select one of 7 landmarks as a missing landmark and then estimate the missing landmark for each skull. Thus, we estimate a total $7 \times 18 = 126$ incomplete configurations. Table 5 gives the values of the MSQRSS for each landmark. We can identify that the results, according to the two algorithms, are quite similar in Table 5. From this table, we can see that most values of MSQRSS are less than 0.05 except for landmark 1 and landmark 3 of males. This means that the estimations of the three-dimensional data are made well.

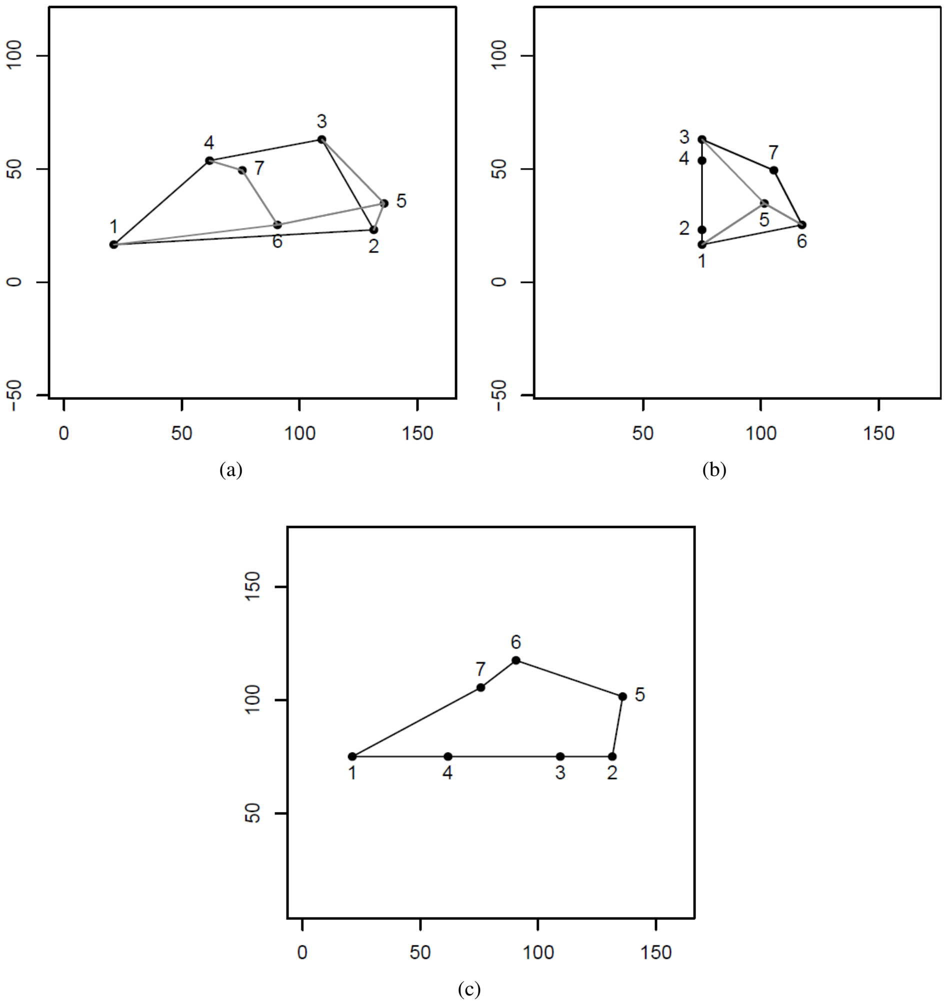


Figure 4: *Seven landmarks on a three-dimensional macaque skull: (a) side view, (b) frontal view and (c) bottom view.*

Table 6 shows the averages of the ratio of the elapsed time for estimating each landmark. From this table we can confirm that the elapsed time of Algorithm II is shorter than that of Algorithm I. The algorithm of Albers and Gower (2010) requires about 190 re-estimates, but about 13 re-estimates are enough.

5. Conclusion

Sometimes, configurations can have one or more missing landmarks, so we need to estimate them. But in statistical shape analysis, typical statistical methods such as expectation-maximization (EM) algorithm or Markov Chain Monte Carlo (MCMC) cannot be used because the data structure in shape analysis is the array form and the shape space is generally considered as a sphere. In Procrustes

Table 5: The MSQRSS of Algorithm I and II for macaque skulls

Delete landmark	Female		Male	
	Algorithm I	Algorithm II	Algorithm I	Algorithm II
1	0.0488	0.0488	0.0585	0.0586
2	0.0326	0.0326	0.0426	0.0426
3	0.0302	0.0302	0.0694	0.0694
4	0.0324	0.0624	0.0388	0.0388
5	0.0342	0.0342	0.0380	0.0380
6	0.0257	0.0257	0.0289	0.0289
7	0.0331	0.0331	0.0400	0.0400

Table 6: The average ratio of the elapsed time of Algorithm I and II for macaque skulls

	Deleted landmark						
	1	2	3	4	5	6	7
Female	16.79	18.37	17.18	17.28	18.75	16.57	16.16
Male	15.59	16.35	16.69	16.10	16.91	17.73	16.80

analysis which is the basis of shape analysis, an algorithm, which can handle missing landmarks is proposed by Albers and Gower (2010). Since GPA algorithm is based on the least square criterion, their algorithm is based on the least square methods. Furthermore, their algorithm can be viewed as the same as a variant of the iterative EM algorithm where M-step, rather than representing a step for maximum likelihood estimation, now stands for the least-square Procrustes problem, while E-step gives expected values for the missing cells. However, their algorithm has the disadvantage of slow convergence to account for the direct transformation of configurations.

In fact, many missing data assignment methods are available in the statistical shape analysis for missing data estimation. It is well known that among mean substitution based on conditional distributions, EM algorithm and multiple regression assignment methods are among the most used missing data assignment methods. Apart from these methods, approaches to estimate missing data by modifying PCA have also been proposed (Nounou *et al*., 2002; Scholz *et al*., 2005; Stacklies *et al*., 2007).

In statistical shape analysis, it is possible to use the geometrical information called shape. The shape is information that remains when location, scale and rotational effects are filtered out from a configuration. Shape can also be measured dissimilarity between or among objects by using Procrustes distance. In this study, we propose a modified algorithm to improve the slow convergence problem. The process of our algorithm is similar to that of Albers and Gower (2010), but we use the shape information for estimating missing landmarks and then measure the shape variability using the sum of squared Procrustes distances. Since it is sufficient to consider only the rotation of the shape in pre-shape space, there is an advantage in convergence speed that is faster than using configuration and the Procrustes sum of squares. Actually, the rate of convergence of our algorithm is definitely shorter than that of Albers and Gower (2010) as seen in our examples.

References

Albers CJ and Gower JC (2010). A general approach to handling missing values in Procrustes analysis, *Advances in Data Analysis and Classification*, **4**, 223–237.

Commandeur JJ (1991). *Matching Configurations*, DSWO Press, Leiden University, Netherlands.

Dryden IL and Mardia KV (1993). Multivariate shape analysis, *Sankhyā: The Indian Journal of Statistics, Series A*, **55**, 460–480.

Dryden IL and Mardia KV (1998). *Statistical Shape Analysis*, John Wiley & Sons, Chichester.

Gower JC (1975). Generalized Procrustes analysis, *Psychometrika*, **40**, 33–50.

Gower JC and Dijksterhuis GB (2004). *Procrustes Problems*, University Press, Oxford.

Nounou MN, Bakshi BR, Goel PK, and Shen X (2002). Bayesian principal component analysis, *Journal of Chemometrics*, **16**, 576–595.

Kendall DG (1984). Shape manifolds, Procrustean metrics, and complex projective spaces, *Bulletin of the London Mathematical Society*, **16**, 81–121.

Kent JT (1994). The complex Bingham distribution and shape analysis, *Journal of Royal Statistical Society, Series B*, **56**, 285–299.

O'Higgins P (1989). A morphometric study of cranial shape in the Hominoidea (Ph.D thesis), University of Leeds, Leeds.

O'Higgins P and Dryden IL (1993). Sexual dimorphism in hominoids: Further studies of craniofacial shape differences in Pan, Gorilla and Pongo, *Journal of Human Evolution*, **24**, 183–205.

Scholz M, Kaplan F, Guy CL, Kopka J, and Selbig J (2005). Non-linear PCA: A missing data approach, *Bioinformatics*, **21**, 3887–3895.

Stacklies W, Redestig H, Scholz M, Walther D, and Selbig J (2007). PcaMethods-A bioconductor package providing PCA methods for incomplete data, *Bioinformatics*, **23**, 1164–1167.

Ten Berge JMF (1977). Orthogonal Procrustes rotation for two or more matrices, *Psychometrika*, **42**, 267–276.

Ten Berge JM, Kiers HA, and Commandeur JJ (1993). Orthogonal procrustes rotation for matrices with missing values, *British Journal of Mathematical and Statistical Psychology*, **46**, 119–134.

Received February 18, 2022; Revised September 02, 2022; Accepted September 15, 2022