

Morphometrics and compositional classes. The study of anthropomorphic sculptures from Teotihuacan (México)

J. BUXEDA i GARRIGÓS¹ and A. VILLALONGA GORDALIZA²

¹Cultura Material i Arqueometria UB (ARQ|UB, GRACPE), Dept. de Prehistòria, Història Antiga i Arqueologia – Universitat de Barcelona, Catalonia, Spain
(jbuxeda@ub.edu)

²Departament d'Art – Universitat Autònoma de Barcelona, Catalonia, Spain

Abstract

Geometric morphometrics is the study of the form and/or shape of objects, or subjects, capturing their geometry. Within the landmark methods developed so far, Euclidean Distance Matrix Analysis (EDMA) is a coordinate-free method that enables the study of form avoiding the nuisance parameters of translation and rotation existing in the landmark coordinates data. However, EDMA exhibits several weaknesses in the study of shape. The present paper links EDMA with compositional data analysis in what is here called Compositional EDMA and illustrates the present approach with the study of anthropomorphic sculptures from the pre-Hispanic site of Teotihuacan (México).

Keywords: geometric morphometrics, EDMA, compositional data analysis, Compositional EDMA, anthropomorphic sculptures, Teotihuacan

1 Morphometrics. A brief introduction

To present, there exist a large literature on morphometrics that provide (1) several definitions for the field itself and (2) for the main concepts involved, as well as (3) several classifications of different approaches to the study of morphometry (see for example Dryden and Mardia, 1998; Lele and Richtsmeier, 2001; Richtsmeier et al., 2002; Verrecchia, 2003; Adams et al., 2007; Slice, 2005, 2007; Lawing and Polly, 2010). In short, it could be said that the form of an object, or a subject, is related to its appearance and structure (Lele and Richtsmeier, 2001, pp. 1), and that morphometry is the measurement of these external and perceptible characteristics, related to the object's appearance as well as to its physical and diachronic constitution (Verrecchia, 2003, pp. 759-760). One important aspect is that even if object's forms, as well as mean forms, could have an orientation in a natural coordinate system, the form itself is invariant under translation, rotation, and, possibly, reflection of the object (Richtsmeier et al., 2002, pp. 72). This is easily illustrated in Figure 1, A. There, the isosceles right triangle (in red) has been first translated and then rotated twice 45° around the centre. The original red triangle and the three new triangles have different orientations, but their form is still the same and has not changed. A further step forward, enables to realize that form consists of size and shape, concepts of an easy intuitive understanding but of a difficult definition (Richtsmeier et al., 2002, pp.

67-68). Nevertheless, as we will see, shape can be considered the result of a central dilation of similar forms and, therefore, shapes can be considered as scale invariants forms that can be related by central dilation after the appropriate similarity transformations (Figure 1, B).

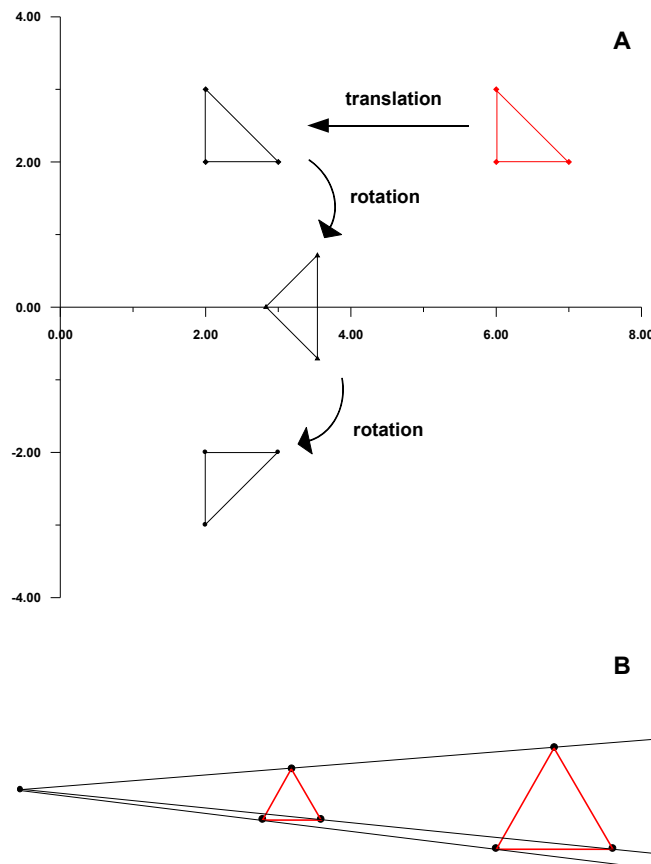


Figure 1: Similarity transformations. A: translation and rotation of an isosceles right triangle. B: central dilation of an equilateral triangle

In morphometric studies, form or shape are approached by using numerical data through mathematical relationships. However, even if this use of numerical data appears very straightforward, several approaches have been developed to present, based on different aspects of form and shape, but also on different ideas on the objects themselves (Figure 2). Summarizing, the oldest approach is the use of unidimensional measures or groups of linear distances or angles. These measures can then be used as such or to calculate ratios. This approach, called traditional or multivariate morphometrics, produces useful results in a large number of situations, and it has also been linked to compositional data analysis. As an example of the latter, it can be seen the study on the use of different proportional systems in Early Christian Churches (Buxeda i Garrigós, 2008), a re-evaluation of a previous study that had revealed the need for the use of the cosine similarity in searching for shape patterns in objects of different size (Gurt Esparraguera and Buxeda i Garrigós, 1996). Nevertheless, traditional morphometrics have proved problematic or limited in different situations because (1) of the need for size correction, was an issue constant debate, (2) the use of non-homologous points, and (3) the possibility to obtain the same distances in different forms (Adams et al., 2007). In short, the use of traditional morphometrics usually implies the loss of the geometry of the object under study.

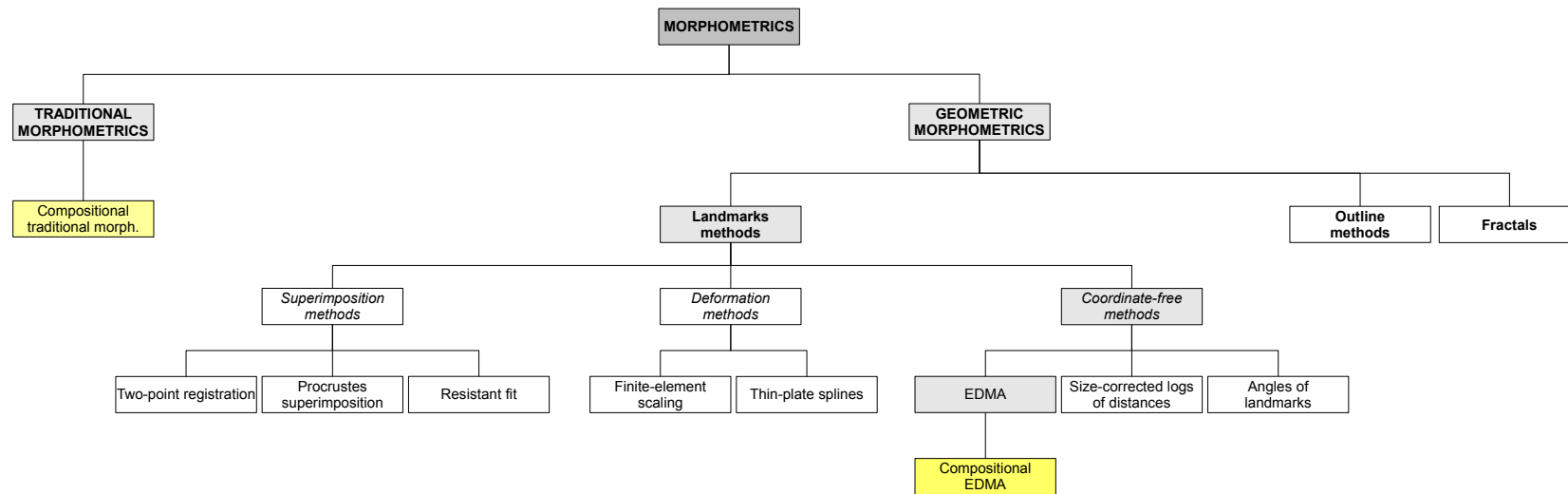


Figure 2: Summary graph of morphometric approaches. In yellow, approaches related to compositional data analysis

To overcome the previous limitations, the so called geometric morphometrics methods (Corti, 1993) were developed from the 1980s in order to capture the geometry of the form under study. The first methods to be used were the outline ones, by reproducing the open and/or closed outlines of the object through several mathematical functions whose coefficients are then used as variables in the statistical analysis (Rohlf, 1990; Dommergues et al., 2007, and references therein). Moreover, fractal dimension measurement have also been used for characterizing the complexity of shape or form (see for example Pérez-Claros et al., 2002; Slice, 2005, pp 38-39). Finally, landmark-based methods are based on precise locations on forms, recorded on two- or three-dimensional coordinates. Thus, the geometry of the form is preserved thanks to the existence of a map of the relative location of those landmarks.

The present paper will explore in some detail the landmark-based methods and especially the Euclidean Distance Matrix Analysis (EDMA) (Lele, 1993; Lele and Richtsmeier, 2001). It will be shown that EDMA is easily linked to Compositional Data analysis (Aitchison, 1986), becoming then a useful method for shape analysis. The anthropomorphic sculptures from the pre-Hispanic site of Teotihuacan (México) will be used as example. These sculptures will be first introduced in the next section.

2 The anthropomorphic sculptures from Teotihuacan

The ancient site of Teotihuacan, located about 50 km to the north-east of México City, was one of the most dynamic and influential cities of the Ancient World (0-600 AD) in pre-Hispanic times. Multi-ethnic city and great handicraft and religious center, the city was planned *ex novo* around the 1st century AD becoming a urban model as a reference for future urban plans. Around 200 AD the city exceeds the 20 km² in extent and two centuries later housed an estimated population of 150,000-200,000 inhabitants (Millon, 1974, pp. 355). The ceremonial center comprises more than 600 pyramids of various kinds and about 2000 complex housing units have been detected so far. Barely 50 out of these 2000 have been partially excavated (Cowgill, 2003, pp. 21).

Focussing on the artistic production at Teotihuacan, it is clear that it involved specialization at various levels. The existence and location of ceramic and obsidian craft workshops testifies its importance and to a which extent the production was under the control of state government institutions. After architecture and mural painting, sculpture is perhaps the craft which less interest has arisen among the research community. Probably the great disparity in material, dimensions, techniques, themes, context, and conditions have hindered its study, although there is a general catalogue on the sculptures in Teotihuacan (Allain, 2000, 2005). In the present study, however, only a more or less homogeneous sculpture assemblage will be considered: the representation of the human figure.

Traditionally, studies of pre-Columbian sculptures tend to be predominantly iconographic, isolating the sample from its archaeological context, both spatially and temporally. There is also a tendency to find analogies to iconographic evidence of later periods, either with other Mexica art works or with evidences in the 16th century AD written sources. As a consequence the studies remain on a very general level and do not deepen enough in the knowledge of these sculptures: “Estos trabajos sólo llegan a redescubrir algunos elementos iconográficos repetidos, o se limitan a identificar tipos de aves, cánidos, felinos, insectos [...] algunos símbolos y, desde luego, los atributos propios de deidades muy conocidas.” (Morelos García, 2002, pp. 30). Then, it is not surprising to read in a catalogue that any Teotihuacan sculpture is ‘Teotihuacan style’, comprising from Tlamimilolpa to Metepec phases, i.e. 200 to 650 AD. This shows, once again, that even if some unification of style might exists, very few sculptures have been dated by archaeological association with ceramic materials.

Thus, the main scope of our research is recovering the link between space, time frame and artistic production (Carrillo and Villalonga Gordaliza, 2007; Villalonga Gordaliza, 2010, 2011). Therefore, all significant information relating the target sculptures to the archaeological projects within whose

frames were recovered has been gathered. The projects and sculptures all together constitute the backbone of our study, providing an informed basis for type classification and for its assignment to temporary phases. Since several of the sculptures proceed from modern well controlled excavations, their analysis also provide reliable data to infer formal features or patterns from them. Along these lines, the spatial-temporal distribution of the sculptures recovered in the city will, in turn, shed light about their distribution and possible function. About the latter, even if some sculptures come from controlled archaeological contexts, their function remains unknown. What or who they represent? Gods, men, or political-religious agents? To determine the function and the role played by these sculptures we consider as essential an approach and analysis based on the space-time context in which the sculptures were conceived. Hence the importance of those sculptures that have survived over time and remained earthed until the day they were excavated, which are unfortunately only the minor part of the existing evidence. On the contrary, most of the known sculptures have been subject of looting, both pre-Hispanic and contemporary, to satisfy the ever increasing demand from museums, collectors and private auctions. These sculptures lie on the shelves of many museums, with no other label that indicates their provenance rather than their alleged affiliation to the city or of belonging to a controversial 'Teotihuacan style'. How can we call it 'Teotihuacan style' when even the frontiers of Teotihuacan itself are well not defined? How much 'Teotihuacan' the sculptures stored in museums are? It is true that there exist previous works which have established some formal traits to Teotihuacan sculptures, but it is also true that they remain in a rather general analysis, not detailed enough (Gamio, 1922; Oropeza, 1968; Nicholson, 2002; de la Fuente, 1985; Sarro, 1988, 1991; Lombardo de Ruiz, 1990; Berlo, 1992; Pasztory, 1993b,a; Berrin and Pasztory, 1993; Lombardo de Ruiz and Nalda, 1996; Michelet and Allain, 2009).

Since we believed there was a possibility that in museums also reliable Teotihuacan sculptures become stored, we expanded the search for anthropomorphic stone sculptures in American, Europe and Oceania museums and museums publications. In our research process we have faced two problems to deal with in the future. On the one hand, we have recovered many fragments of what were once complete sculptures. In some cases the destruction corresponds to the different postdepositional processes undergone during the formation of the final archaeological record. In other cases we are facing human activities intentionally causing the destruction of pre-Hispanic rituals, as in the centuries of conquest. On the other hand, the review of some of the sculptures found in museums and private collections, have revealed the world of fakes, the existence of possible imitations that distort any study of such remains.

In a first approach, a group of 39 sculptures representing males with headdress has been studied (Table 1), while the sample is currently increasing to accommodate a much larger number of individuals. As can be seen, most of them do not have known archaeological context of recovery, but a recovery at Teotihuacan is postulated. A different situation is encountered with those individuals that were unearthed at the Feathered Serpent Pyramid (Temple of Quetzalcoatl). This pyramid was built in Miccaotli or Early Tlamimilolpa phases (150-250 AD), in a single episode, and its construction involved the sacrifice of around 200 individuals with their military gear, when corresponding, and with other material elements. A especial case is the burial 14, with some 20 males with very rich offerings and, among them, several of the sculptures here studied (Cabrera Castro et al., 1991). In a later study, Sugiyama (2005) dated the pyramid construction to around 210 AD, after several ¹⁴C dates. In the same work, this author studied the anthropomorphic sculptures that "[...] were apparently scattered on the bodies of the sacrificed victims." (Sugiyama, 2005, pp. 149), and concluded that they exhibit different forms, measures, and proportions. After classifying them in 8 groups (A to H) the author concluded that "Their typological differences and similarities may have ritual significance and/or social implications; the differences also would have reflected variables not directly related to the mortuary program such as temporal variation in productions or different craftsmen. Their variety may even indicate that each piece represented a specific individual or group." (Sugiyama, 2005, pp. 149-150).

Individual	Original location	Present location	Chronology
1	?	MNAM, México	?
2	?	MNAM, México	?
9	?	Musée du Quai Branly, Paris	150-650 AD
13	?	Musée du Quai Branly, Paris	150-650 AD
35	?	Museum für Völkerkunde, Vienna	200-600 AD
39	?	Saint Louis Art Museum	200-600 AD
40	?	Saint Louis Art Museum	200-600 AD
42	?	Private Collection	100 BC-400 AD
47	?	Museum für Völkerkunde, Vienna	?
51	?	Art Institute of Chicago	400 AD
53	?	Museum of Fine Arts, Houston	?
56	?	Natural History Museum, Los Angeles	400-600 AD
57	?	Natural History Museum, Los Angeles	400-600 AD
60	?	Denver Art Museum	200-600 AD
64	?	Denver Art Museum	400-700 AD
74	?	Überseemuseum Bremen	?
76	?	Los Angeles County Museum of Art	300-600 AD
100	?	National Museum of the American Indian, Washington	1-600 AD
113	FSP?	Museo de Sitio, ZAT	?
115	?	Museo de Sitio, ZAT	1-600 AD
127	?	MNAM, México	?
128	?	MNAM, México	?
129	?	MNAM, México	?
132	N3W1Zona6 PT60-62	Acervo de la ZAT	?
191	?	Musées Royaux d'Art et Histoire, Bruxelles	150-200 AD
203	?	Israel Museum, Jerusalem	?
204	?	Israel Museum, Jerusalem	?
228	FSP 1939Templo Quetzalcoatl, Offering 1	TMPL, San Juan Teotihuacan	150-250 AD
229	FSP 1939Templo Quetzalcoatl, Stairs Offering	TMPL, San Juan Teotihuacan	150-200 AD
M91750	?	Bóveda MNAM, México	?
9C-2493	FSP PTQ89, Burial 14	TMPL, San Juan Teotihuacan	ca. 210 AD
9C-3295	FSP PTQ, Burial 14	TMPL, San Juan Teotihuacan	ca. 210 AD
M91751	?	Bóveda MNAM, México	?
M91739	?	Bóveda MNAM, México	?
9C-2943	FSP PTQ89, Burial 14	TMPL, San Juan Teotihuacan	ca. 210 AD
9C-2858	FSP PTQ89, Burial 14	TMPL, San Juan Teotihuacan	ca. 210 AD
9C-3368	FSP PTQ89, Burial 14	TMPL, San Juan Teotihuacan	ca. 210 AD
9C-2986	FSP PTQ89, Burial 14	TMPL, San Juan Teotihuacan	ca. 210 AD
E-16	FSP PTQ89, Area B, Well 3	TMPL, San Juan Teotihuacan	ca. 210 AD

Table 1: Anthropomorphic sculptures included in the present sample. FSP: Feathered Serpent Pyramid (Temple of Quetzalcoatl), Teotihuacan. MNAM: Museo Nacional de Antropología de México. TMPL: Teotihuacan Mapping Project Laboratory. ZAT: Archaeological Area of Teotihuacan. Chronologies have been provided by museums. For sculptures from FSP PTQ89, Burial 14 chronology is taken from Sugiyama (2005)

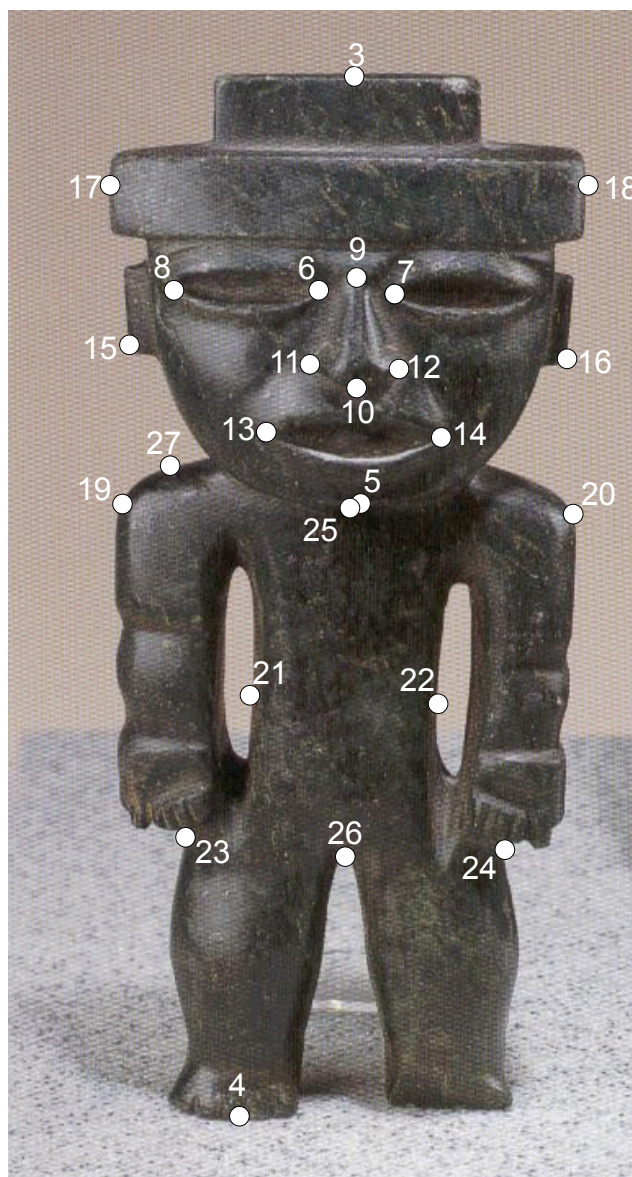


Figure 3: Anthropomorphic sculpture with indication of the 25 landmarks initially recorded

The coordinates of 25 landmarks (Figure 3), or fixed points of precise location, were recorded for the 39 individuals studied. These two-dimensional coordinates were recorded from pictures of the actual sculptures with the help of image processing software, and not from the sculptures themselves. Such decision was needed because for some of the individuals, pictures were the only available record at our disposal. Nevertheless, since all these sculptures are not really developed in three dimensions and they appear rather as a bas relief, two-dimensional coordinates seem appropriate for the present study. It must be noticed, however, that even if the use of image software can give a higher precision to the recording of coordinates, the use of pictures could might introduce distortions because of deformation of the pictures, but also because the angle at which the picture was taken. Even if we assume that these factors should be of very limited impact on the data recorded, we have no conducted any measurement error study.

3 EDMA and Compositional EDMA

As it will be shown, Euclidean Distance Matrix Analysis (EDMA) is a landmark-based method that exhibit advantages in front of superimposition and deformation methods. Furthermore, it is easily linked to Compositional Data Analysis. In this section we shall present the main features of EDMA, and we will make clear the connexion with Compositional Data Analysis. In this presentation we will make a quite extensive use of the the studies by Lele and co-workers, who have develop EDMA (see especially Lele, 1993; Lele and Richtsmeier, 2001; Lele and McCulloch, 2002; Richtsmeier et al., 2002, 2005).

3.1 Landmarks, coordinates, and invariance

Landmarks, as stated before, are precise locations on forms, recorded on two- or three-dimensional coordinates. A more formal description is given by Dryden and Mardia (1998, pp. 3) as “[...] a point of correspondence on each object that matches between and within populations.”. Since geometric morphometrics is mainly related to biological studies, several classifications have been proposed, based on biological criteria, for different types of landmarks. Thus, Bookstein (1991) distinguishes among Type I, Type II, and Type III landmarks, respectively corresponding to discrete juxtaposition of tissues, curvature maxima associated to local structures, and extremal points. Dryden and Mardia (1998, pp. 3-6) differentiate (1) anatomical landmarks, points biological meaningful, (2) mathematical landmarks, points located according to some mathematical or geometrical property, (3) pseudo-landmarks, or constructed points on an organism, and (4) semi-landmarks. Moreover, these authors relate their landmark types with those established by Bookstein. They also differentiate between labelled and unlabelled landmarks. Finally, Lele and Richtsmeier (2001, pp. 24-27) classify landmarks as (1) homologous landmarks, similar features on phylogenetic basis, (2) structurally corresponding landmarks, (3) functionally corresponding landmarks, and (4) developmentally corresponding landmarks. For our purposes, we shall be talking about landmarks, without any further implication that regarding points of precise locations in all studied individuals. Therefore, the landmarks used in this study can be considered as *corresponding* landmarks (Richtsmeier et al., 2002).

An important feature of landmarks is that they preserve the geometry of the form thanks to the existence of a map of their relative location, as given by the coordinates. However (see especially Richtsmeier et al., 2002), if the landmarks have a natural coordinate system, these natural coordinates are lost and cannot be known, since the objects under study could have suffered translations and/or rotations that cannot be identified. Thus, landmarks are not recorded in their natural coordinates, but in arbitrary coordinates. And those coordinates should be further transformed to some coordinate system useful for the statistical analysis. At his point, superimposition and deformation methods use different rules to take back all objects under study to a common coordinate system. In doing so, these methods make arbitrary choices on how to achieve a common coordinate system that affect the results that will be obtained. Moreover, those methods do not really remove translation and rotation, which are nuisance parameters in landmark analysis. Because of that, “[...] the popular methods of superimposition and deformation base inferences on nonidentifiable parameters.” (Lele and McCulloch, 2002, pp. 805).

Contrariwise, EDMA method effectively removes the nuisance parameters by using a maximal invariant, under translation, rotation, and possibly reflection, of this this precise locations of landmarks, i.e. all possible linear distances among pairs of landmarks (see especially Lele, 1993). Thus, EDMA is an invariant under nuisance parameters coordinate-free method, that preserves the map of relative location, and amending therefore to base inferences on identifiable parameters of interest. In practical terms, the use of all possible linear distances among pairs of landmarks will led to produce the appearance of two- or three-dimensional forms of lines connecting landmarks. If we look back to Figure 1, A and B, we can see that three landmarks produce the appearance of triangles (as it is

indicated by the points, the initial landmarks, in the vertices of each triangle). Finally, it must be stressed that if a form is initially orientated in the natural coordinate system, the form invariant under translation, rotation, and reflection, i.e. the infinite equal forms of all possible orientations, occupy an orbit defined by this particular form.

3.2 Perturbation model, morphometric spaces, and parameters of interest

The usual model for the landmarks variability is the so called general or Gaussian perturbation model (Goodall, 1991; Lele, 1993; Lele and McCulloch, 2002). In that model, the mean form of a K point configuration in L -dimensional Euclidean space of a population is represented by the $K \times L$ matrix M . Each individual of this population deviates from M by the $K \times L$ matrix E_i of errors. E_i is assumed to be Gaussian, $E_i \sim N(0, \Sigma_K, \Sigma_L)$, where Σ_K describes the covariances between elements within the same same column of E and Σ_L describes the covariances within the rows of E_i . Producing a unique vector from the E_i matrix, $\text{vec}(E_i)$, one obtain that $\text{var}(\text{vec}(E^T)) = \Sigma_K \otimes \Sigma_L$. Thus variability, in the natural space, has the amount of variability at each landmark, the shape of this variability (related to the possible correlations between the coordinates of each landmark), and the possible correlations between different landmarks. Since the natural coordinates are unknowable because of the nuisance parameters of translation and rotation, the landmark coordinate matrix of the i th individual is then

$$X_i = (M + E_i)\Gamma_i + 1t_i^T, \quad (1)$$

where Γ_i is the orthogonal matrix of rotation, and possibly reflection, $\mathbf{1}$ is a $K \times 1$ matrix of 1s, and t_i^T is the $L \times 1$ matrix of translation. Then the random matrices X_i follow:

$$X_i \sim N(M\Gamma_i + 1t_i^T, \Sigma_K, \Gamma_i^T \Sigma_L \Gamma_i). \quad (2)$$

Based on such perturbation model, a further step forward is the definition of different 'spaces' for morphometric analysis. Among them the most widely used are (1) Kendall's shape space, i.e. an sphere, (2) Kent's tangent shape space, a plane tangent to the sphere, and (3) form space (Kendall, 1984; Dryden and Mardia, 1998; Lele and Richtsmeier, 2001). The latter is the space developed in EDMA (Lele and Richtsmeier, 2001). The idea behind the form space is quite simple and intuitive. Thus, if the i th individual has a K point configuration in L -dimensional Euclidean space giving a $K \times L$ natural space coordinate matrix X_i , once its landmarks have been recorded in an arbitrary coordinate system, it has a $K \times L$ landmark coordinate matrix A_i . Then, the set of all possible linear euclidean distances among its pairs of landmarks gives its $K \times K$ form matrix

$$FM_i = \begin{pmatrix} 0 & d_{12} & \cdots & d_{1k} \\ d_{21} & 0 & \cdots & d_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ d_{k1} & d_{k2} & \cdots & 0 \end{pmatrix}$$

Since in this form matrix the number of unique pairs of linear euclidean distances is D , with $D = K(K-1)/2$, the vector of these D distances, $\text{vec}(d-(ij))$, configures a point in the form space, a subset of the positive orthant of the $D = K(K-1)/2$ Euclidean space, which uniquely identifies an orbit of a given form (see especially Lele and Richtsmeier, 2001, pp. 146 ff.). Thus, for example, a 3 landmarks configuration gives a 3-dimensional ($D = 3(3-1)/2$) form space. As can be seen in Figure 4 the 3-dimensional form space, the subset of the 3-dimensional euclidean space in yellowish lines, will contain all points identifying forms of all possible triangles.

Finally, it must be noticed that the use of the form matrices and the form space will enable, by the method of moments, the identification of several aspects of the parameters of interest (M, Σ_K, Σ_L), i.e. (1) the mean form M , up to its orbit, (2) a singular version of Σ_K ($\Sigma_K^* = L\Sigma_K L^T$, where L is a centering matrix), and (3) eigenvalues of Σ_L . It is important to notice, however, that the singular matrix Σ_K^* . The

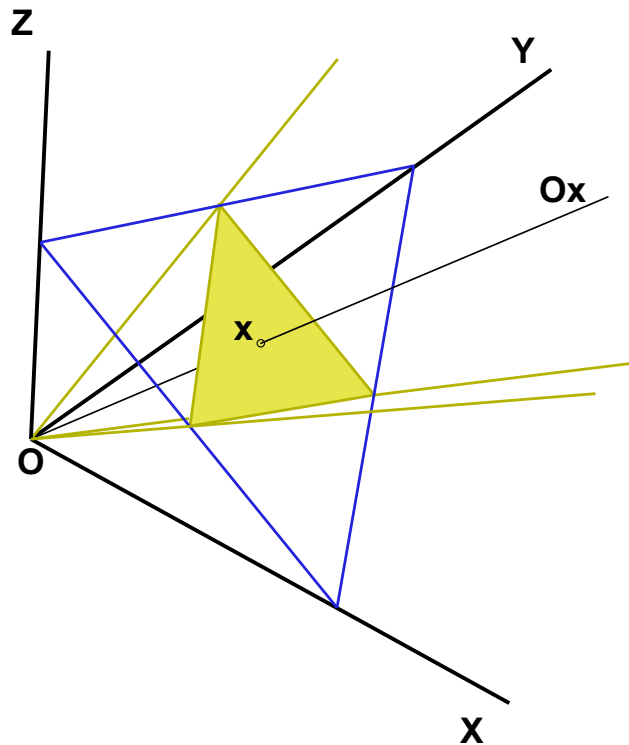


Figure 4: Form space for three-landmark objects (triangles). This drawing is partially inspired in the one at Lele and Richtsmeier (2001, pp. 147)

usefulness of this estimate is that it can be used in confidence intervals and hypothesis tests without any effect on the centering matrix used in its calculation (Lele, 1993; Lele and McCulloch, 2002; Richtsmeier et al., 2005).

3.3 Form and shape, size, and projective points and Compositional EDMA

One of the main drawbacks of EDMA, according to its opponents, is that it is a method for the study of form, not really for the study of shape (see discussion in Richtsmeier et al., 2002, pp. 88). Form, in EDMA, is studied either by arithmetic form difference (for example, between the i th and j th individuals is given by $AFDM_{ji} = FM_j - FM_i$) or by relative form difference ($FDM_{ji} = FM_j / FM_i$, done element-wise and using $0/0 = 0$ by definition).

Dealing with shape in EDMA starts from the view that form should be scaled by size, but that there is no clear definition of size. On the contrary, there are many surrogate measures of size that will result in different size-corrected shapes (see discussion in Richtsmeier et al., 2002, pp. 67 ff.). The origin of this view is the discussion about ratios and size, first clarified by Mosimann (1970), and further discussed by other authors (Atchley et al., 1976; Packard and Boardman, 1999). In EDMA, it is proposed to use the geometric mean as a measure of size ($S_i = (\prod FM_i)^{1/D}$). Then scaling by this size, one gets the shape matrix for the i th individual ($SM_i = FM_i / S_i$). Then one can calculate the shape difference between two individuals just as an arithmetic shape difference.

Even though this possibility for 'shape' analysis exists in EDMA, if we look at Figure 4, we could see that all equilateral triangles, all orbits of equilateral triangles forms, may lay on the ray Ox , a projective point, i.e. a $d+1$ -dimensional vector of coordinates, or coordinate vector, $\mathbf{x} = [x_1, \dots, x_{d+1}]$ (enclosed within brackets), where at least one of the x_i is nonzero; x_i are called homogeneous, or projective, coordinates. These homogeneous coordinates are, then, the components of a $d+1$ -part composition (where $D = d+1$) (Buxeda i Garrigós, 2008). The observation of this fact leads naturally to the realization that scale invariant forms are compositional or equivalence classes $[\mathbf{x}] = \{t\mathbf{x} \mid t \in \mathbb{R}^{d+1}, t \neq 0\}$,

leading to Compositional EDMA.

To the best of our knowledge, this approach has not been really explored, neither used in the past. Nevertheless, the idea that the interpoint distance of landmarks could be use as input data for Compositional Data Analysis has been already suggested by Bacon-Shone (1991). Moreover, also Dryden and Mardia (1998, pp. 27) suggested the possibility to link geometric porphometrics with Compositional Data Analysis. Even so, this link was limited to the angles of triangles, because of their closure operation to 180° .

According to what has been shown so far, shape can be defined as the scale invariant approach to those compositional or equivalence classes (or projective points) existing in the form space defined by all unique pairs of linear euclidean distances among landmarks. This form space is a subset of the positive orthant of the euclidean space. And, contrariwise to the discussion about scaling by size, CLR transformation is a maximal invariant that naturally introduces Aitchison metric in the study of shape (Aitchison, 1986, 1992, 2005).

If we look at compositional classes (Figure 5) we can see that there are several properties that can be divided in structural and non-structural (Magurran, 2004; Aitchison, 1986). Structural properties include (1) the dimension of the system, the $d+1$ -dimensions of the class, (2) its richness, the number of dimensions with values over zero, and (3) the evenness, the distribution of the values in all dimensions. Diversity, thus, increases by higher richness, but also by more even distributions. On the other hand, non-structural properties include (1) the parts, the labelling of the dimensions, (2) the components, the numerical values of each dimension, and (3) the size, the addition of all values of all dimensions. The relation between components and size will determine the relative values of the components and, therefore, their class.

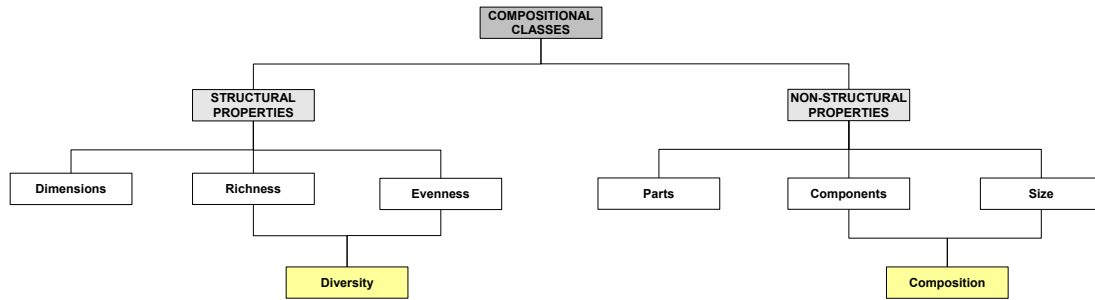


Figure 5: Structural and non-structural properties of compositional classes

Evenness is thus a structural characteristic of compositional classes and it is measured by information entropy (Shannon, 1948). This value is usually expressed in bits, since calculus is made by base 3 logarithms, according to the following equation:

$$H = -k \sum_{i=1}^D p_i \log p_i. \quad (3)$$

This equation can be expressed to accommodate not just probabilities, but also actual components (n_i) and size (N) of each composition:

$$H_f = -k \sum_{i=1}^D \frac{n_i}{N} \log \frac{n_i}{N}, \quad (4)$$

and, finally, given that for a given system with $d+1$ -dimensions the maximum value for H is $\log D$, relative information entropy is given by:

$$H_{rel} = \frac{H_f}{\log D}. \quad (5)$$

As can be seen in Figure 6, H_f enables to quantify information entropy for compositional classes in the form space, and can also be graphically displayed, for a 2+1-dimensional space, over the 2-dimensional simplex. Thus, each compositional class has its own value in information entropy, indicating their lower or higher evenness.

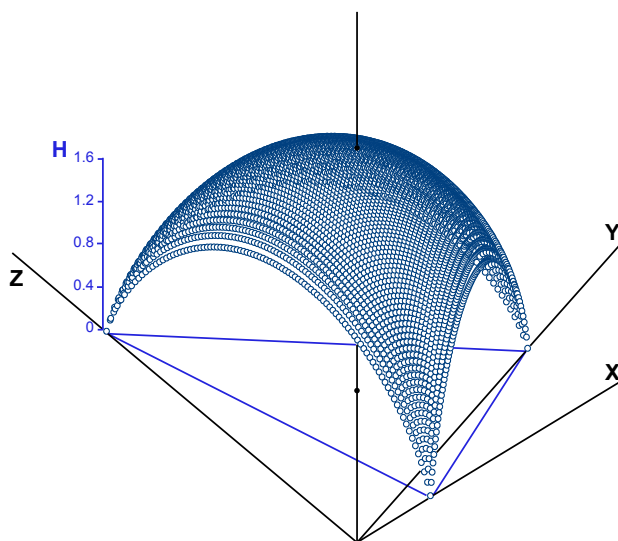


Figure 6: Information entropy as a structural property of evenness of compositional classes (projective points)

Once all these points have been considered, it is important to notice that the only form space that can achieve the highest possible entropy for a given $d+1$ -dimensional system is that of three-landmarks objects, i.e. the case of triangles, whether the equilateral triangle shows the highest information entropy possible, being placed at the center of the system (see Figure 4, in blue the 2-dimensional simplex, in yellowish colors the area in the simplex represented in the resulting form space). If we extend these observations to other polygons, from their regular form to their collapse in a pathological form, we will observe that any other form space cannot include the higher evenness compositional classes (Figure 7). Moreover, the subset of the corresponding euclidean space they include varies in significant terms.

To finish, it is important to notice that once has been demonstrated that Compositional EDMA effectively links EDMA with Compositional Data Analysis for shape analysis, all developments already existing for Compositional Data Analysis can be applied to the study of shape in geometric morphometrics. Especially, perturbation, as it is known in Compositional Data Analysis, can be easily employed to model shape change. In doing so, it is necessary to review a final point. If we are studying N individuals (like the present 39 anthropomorphic sculptures), in EDMA at the beginning one is dealing with $N K \times L$ landmark coordinate matrices. In a second step, these landmark coordinate matrices are transformed in $N K \times K$ form matrices. Then, in a third step, what we have to do is to transform these N form matrices in one single $N \times D$ matrix X , with every individual in one row, expressed by the D unique pair-wise landmark linear euclidean distances. This matrix, can then be used to model perturbation, as shape change, according to:

$$X_n = X_0 \circ U, \quad (6)$$

where U is the $N \times D$ matrix of perturbing vectors. Thus, perturbation as shape change is modelled in the pair-wise landmarks linear euclidean distances.

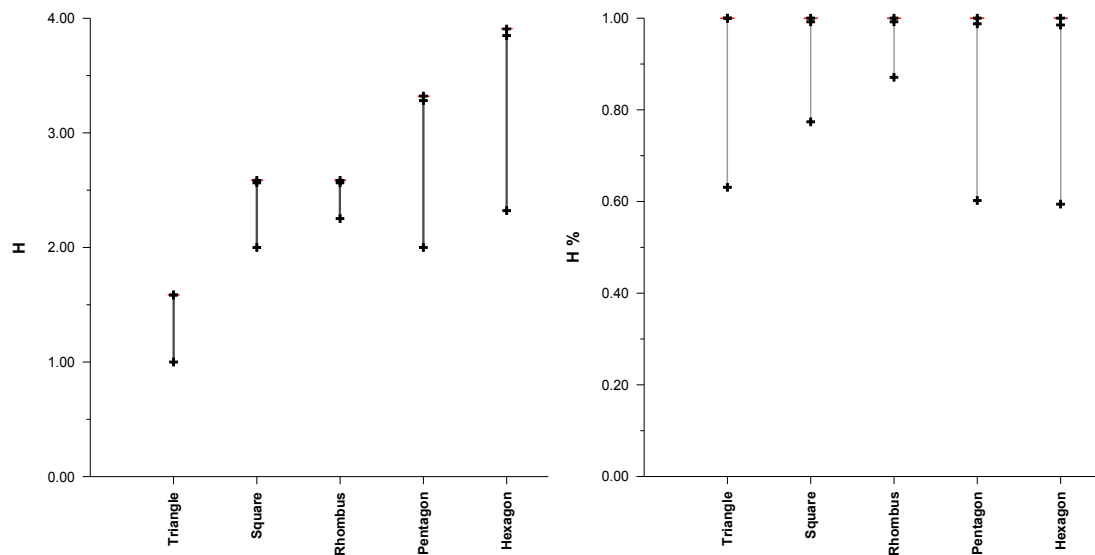


Figure 7: Information entropies, H , for several polygons, from their regular shape to their pathological collapsed one. Left: in bits. Right: in % of total information entropy of the system. In both cases, the red line points the maximum possible entropy for the euclidean space where the form space is a subset from

3.4 The anthropomorphic sculptures

As an example of application, we give in this section some preliminary results on the anthropomorphic sculptures from Teotihuacan (México). Since this is a work still ongoing and the sample is still increasing with new individuals, the results cannot be considered definitive. On the contrary, their main value is to illustrate the preceding discussion and to prove that Compositional EDMA is an useful tool to shed light in studying shape and, therefore, proportions systems in archaeological materials.

As it has been pointed in the previous sections, the initial landmark coordinates matrices have been transformed first into form matrices and then into a single $N \times D$ form matrix X . Since the number of landmarks used at the beginning was quite large (25), the matrix X is a 39×300 matrix. As it is clear, the number of dimensions is clearly too high to enable any analysis. This, the dramatic increase of dimensions in the form space as the number of landmarks increase, is clearly one of the main problems, even drawbacks, in EDMA and, therefore, in Compositional EDMA. In order to overcome this problem, one should conduct different partial analysis addressing specific questions that should then be combined in drawing the final conclusions.

As a first analysis, then, we have decided to retain for analysis landmarks 3, 4, 5, 9, 17, 19, 23, and 26 (Figure 8) addressing the question of the general shape. The matrix is then a 39×28 set that can be studied.

The variation matrix (Aitchison, 1986; Buxeda i Garrigós, 1999) enables to quantify the variability existing in a data set (Buxeda i Garrigós and Kilikoglou, 2003). Moreover, the totals of each column can be used to estimate the contribution to the total variation by each part. A graphical representation can then be made with an evenness plot, giving in the x-axis the different parts in decreasing rank order of variability, while in the y-axis are represented the components. In the present case (Figure 9) it can be seen that few linear distances, 8 out of 28, are responsible for most of the total variation.

Cluster analysis on the CLR transformed data, the shape data, was then performed, with SPlus (MathSoft, 1999), using the square euclidean distance and the centroidal agglomerative algorithm. The study of the dendrogram (Figure 10) shows the existence of two very distinct individuals, F127 and F204. The remainder 37 individuals, however, still exhibit important differences. Nevertheless, tentatively, two groups have been defined, A and B. At this point, it is important to highlight that both



Figure 8: Anthropomorphic sculpture with indication of the 8 landmarks, in yellow, retained for the analysis of the general shape

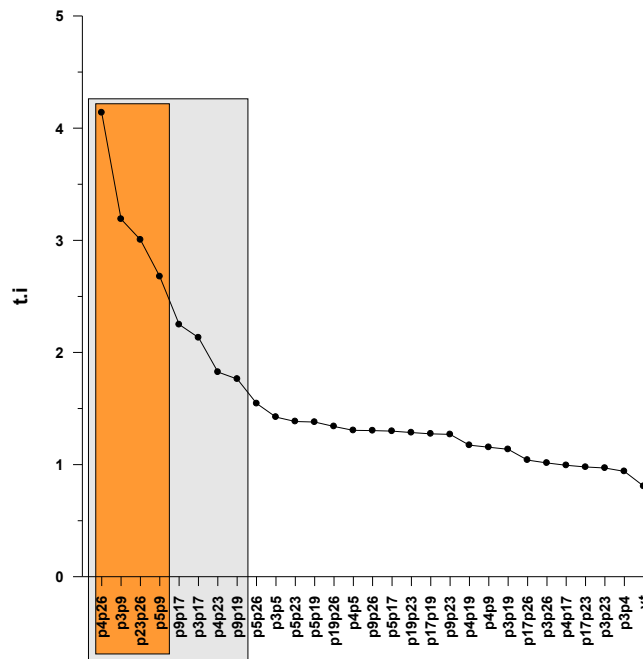


Figure 9: Evenness graph of the variation matrix for 8 landmarks in the general shape analysis

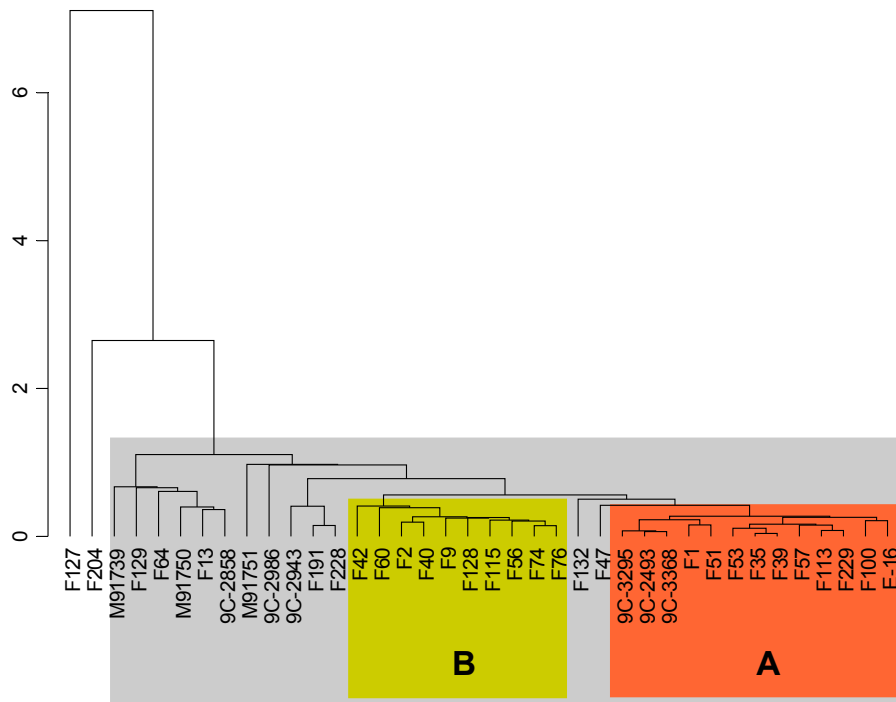


Figure 10: Dendrogram from cluster analysis on 8 landmarks in the general shape analysis

groups include individuals of very different sizes. Group A includes individuals from ca. 3.5 cm to over 10, while group B includes individuals from 13 cm to over 30 cm. This result clearly indicates the adequacy of Compositional EDMA to work with shape in a scale invariant approach.

Finally, a second analysis has been conducted on landmarks 5, 6, 8, 9, 10, 11, and 13 (Figure 11), i.e. a 39×21 matrix, in order to address the question of the face shape.

The present evenness graph (Figure 12) shows a much higher total variation than in the previous case. This is a very important point, since it must be noticed that only 7 landmarks are in use in the present analysis. Moreover, the evenness graph also exhibits the influence of few linear distances in the total variation, but the graph shows a continuous decreasing in variation which is quite different from the previous analysis.

Cluster analysis on the CLR transformed data, the shape data, was then performed, with SPlus, using the square euclidean distance and the centroid agglomerative algorithm. The study of the dendrogram (Figure 13) shows the existence of four very distinct individuals, F127, 9C-2858, F128, and F229, while the remainder 35 individuals seem quite similar. In any case, the existence of different subgroups is still under consideration. If so, it should be noticed that all individuals whose provenance is at the Feathered Serpent Pyramid are located in very different places of the dendrogram. Since all these individuals are of a very short size, around (3 to 4) cm, this could be an indication of the limitations encountered in small objects. These limitations are not necessarily from the landmarks coordinates recording methods, but for the ancient craftsmen themselves most probably in producing such small objects.

As it has been seen, then, Compositional EDMA is a powerful tool for shape, as scale invariant form, analysis. In the present case study, the results are necessarily preliminary. Even though, they are also promising.

A final point that must be highlighted is that once the mean shape form for a groups has been defined, it is possible in Compositional EDMA, as it is also in EDMA. Lets suppose that

$$\underline{\mathbf{M}}^T = (d_{12}, d_{12}, \dots, d_{1D}, d_{23}, d_{24}, \dots, d_{2D}, \dots, d_{(D-1)D})$$

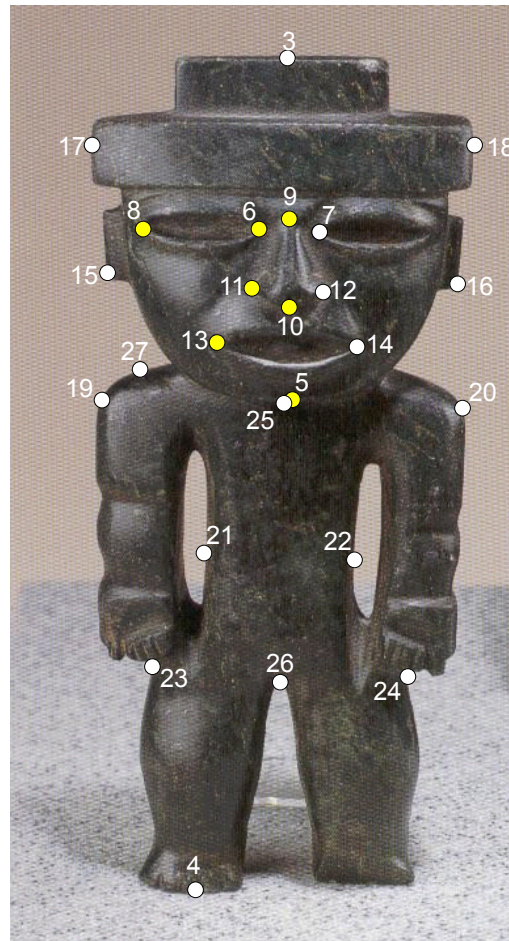


Figure 11: Anthropomorphic sculpture with indication of the 7 landmarks, in yellow, retained for the analysis of the face shape

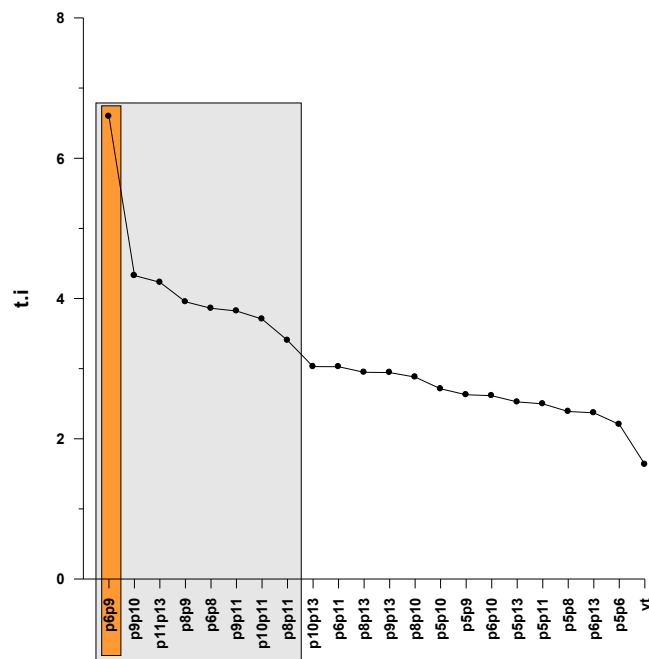


Figure 12: Evenness graph of the variation matrix for 7 landmarks in the face shape analysis

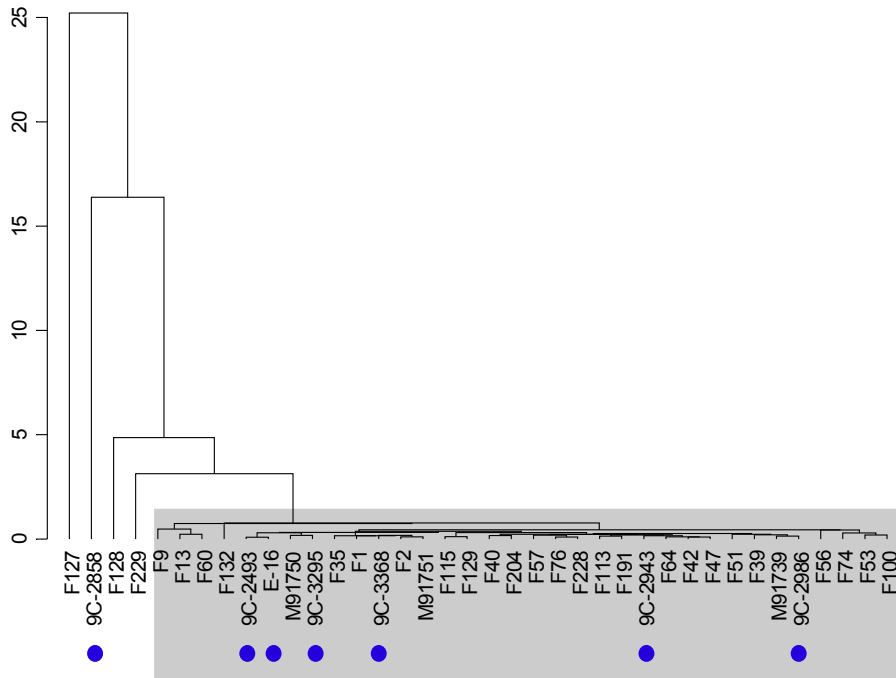


Figure 13: Dendrogram from cluster analysis on 7 landmarks in the face shape analysis. Blue circles: sculptures from the Feathered Serpent Pyramid

is the vector of means for one group, in shape (compositional) analysis, i.e. in CLR transformed data. Then, the only thing to do is to transform this vector by taking e to the power of the different mean components. The resulting vector

$$\underline{\exp M^T} = (e^{d_{12}}, e^{d_{12}}, \dots, e^{d_{1D}}, e^{d_{23}}, e^{d_{24}}, \dots, e^{d_{2D}}, \dots, e^{d_{(D-1)D}})$$

can then be converted back to a matrix of mean distances (see especially Lele and Richtsmeier, 2001, pp. 207-209). Then, the values of such matrix should be squared and, as it is standard in multidimensional scaling, the matrix must be centered. Finally, the spectral decomposition of such centered matrix will enable to obtain the non-zero eigenvalues and eigenvectors that will be used to construct the matrix of landmark representation of the mean shape of the group.

4 Conclusions

The present paper has presented Compositional EDMA as a natural extension of EDMA and Compositional Data Analysis. The realization that orbits in form analysis regarding similar objects are scale invariant projective points, i.e. compositional or equivalence classes, has enabled an easy and robust analysis of shape in coordinate-free geometric morphometrics. All advances in Compositional Data Analysis can then be easily applied to geometric morphometrics.

Regarding the archaeological case study, even in the results are preliminary, they are nonetheless encouraging. It has been shown that one of the main problems is the large number of form space

dimension arising even for moderate number of landmarks. In such situation, several distinct analysis focussing in specific questions should be addressed. In the present case, the anthropomorphic sculptures from Teotihuacan seem to exhibit different shapes, that could be related to different systems of proportions, i.e. to different traditions. Compositional EDMA has also proved useful in analysing objects of quite different shapes. Only the extension of the present work, however, will shed light on the present preliminary shadows.

Acknowledgements

The authors are indebted to Dra. Diana Magaloni Kerpel, Dra. Martha Carmona Macías and Mtro. Enrique Soruco (Museo Nacional de Antropología de México), Lic. Alejandro Sarabia (Zona Arqueológica de Teotihuacan), Mtro. Rubén Cabrera Castro (Proyetco Teotihuacan 80-82 and Proyecto Templo de Quetzalcoatl 88-89), Dr. George L. Cowgill (Emeritus at Arizona State University), Dra. Oralia Cabrera (ASU-Managed Teotihuacan Archaeological Research Center), Dra. María Teresa Uriarte Castañeda (IIE-UNAM), and Dr. Eric Taladoire (Université Paris 1 Panthéon-Sorbonne) for enabling the access to the archaeological materials and for fruitful discussion and guidance on different issues of this research. The development of Compositional EDMA approach is included in the project *Tecnolonial-Technological impact in the colonial New World. Acculturation in pottery archaeology and archaeometry* (HAR2008-02834/HIST) funded by the Ministerio de Ciencia e Innovación (Spain).

References

- Adams, D. C., Rohlf, F. J., and Slice, D. E. (2007). Geometric morphometrics: ten years of progress following the 'revolution'. *Italian Journal of Zoology*, 71:5–16.
- Aitchison, J. (1986). *The Statistical Analysis of Compositional Data*. Chapman and Hall, London.
- Aitchison, J. (1992). On Criteria for Measures of Compositional Difference. *Mathematical Geology*, 24:365–379.
- Aitchison, J. (2005). *A concise guide to compositional data analysis. 2nd Compositional Data Analysis Workshop – CoDaWork'05*. Universitat de Girona, Girona.
- Allain, A. (2000). *Inventaire de la sculpture en ronde-bosse à Teotihuacan*. Memoire de DEA en Archéologie Précolombienne, Université de Paris I-Panthéon, Sorbonne, Paris.
- Allain, A. (2005). *La sculpture dans la civilisation de Teotihuacan*. Thèse pour obtenir le grade de docteur, Université de Paris I-Panthéon, Sorbonne, Paris.
- Atchley, W. R., Gaskins, C. T., and Anderson, D. (1976). Statistical Properties of Ratios. I. Empirical Results. *Systematic Zoology*, 25(2):137–148.
- Bacon-Shone, J. (1991). Discussion of the paper by Goodall. *Journal of the Royal Statistical Society. Series B (Methodological)*, 53(2):332.
- Berlo, J. C. (1992). Icons and Ideologies at Teotihuacan: The Great Goddess Reconsidered. In Berlo, J. C., editor, *Art, Ideology and the City of Teotihuacan*, pages 129–168. Dumbarton Oaks, Washington, D.C.
- Berrin, K. and Pasztory, E. (1993). *Teotihuacan. Art from the city of gods*. Thames & Hudson, London.

- Bookstein, F. L. (1991). *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press, Cambridge.
- Buxeda i Garrigós, J. (1999). Alteration and Contamination of Archaeological Ceramics: the Perturbation Problem. *Journal of Archaeological Science*, 26:295–313.
- Buxeda i Garrigós, J. (2008). Revisiting the compositional data. Some fundamental questions and new prospects in Archaeometry and Archaeology. In Daunis-i-Estadella, J. and Martín-Fernández, J., editors, *Proceedings of CODAWORK'08, The 3rd Compositional Data Analysis Workshop, May 27-30*, pages 1–18. University of Girona, Girona.
- Buxeda i Garrigós, J. and Kilikoglou, V. (2003). Total variation as a measure of variability in chemical data sets. In van Zelst, L., editor, *Patterns and Process. A Festschrift in honor of Dr. Edward V. Sayre*, pages 185–198. Smithsonian Center for Materials Research and Education, Suitland, Maryland.
- Cabrera Castro, R., Sugiyama, S., and Cowgill, G. L. (1991). The Templo de Quetzalcoatl Project at Teotihuacan: A Preliminary Report. *Ancient Mesoamerica*, 2(2):77–92.
- Carrillo, E. and Villalonga Gordaliza, A. (2007). Análisis cerámico del conjunto de los Glifos del Barrio de La Ventilla (2004) de Teotihuacan. In Solanilla Demestre, V., editor, *Arte y Arqueología en Teotihuacan: nuevos trabajos. Actas*, pages 48–57. Grup d'Estudis Precolombins, Departament d'Art, Barcelona.
- Corti, M. (1993). Geometric morphometrics: an extension of the revolution. *Trends in Ecology and Evolution*, 8:302–303.
- Cowgill, G. L. (2003). Teotihuacan. Ciudad de misterios. *Arqueología Mexicana*, XI(64):20–27.
- de la Fuente, B. (1985). *Peldaños en la conciencia. Rostros en la plástica prehispánica*, volume 39 of *Colección de arte*. UNAM, México D.F.
- Dommergues, C. H., Dommergues, J.-L., and Verrecchia, E. P. (2007). The Discrete Cosine Transform, a Fourier-related Method for Morphometric Analysis of Open Contours. *Mathematical Geology*, 39:749–763.
- Dryden, I. L. and Mardia, K. V. (1998). *Statistical Shape Analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons, Chichester.
- Gamio, M. (1922). *La población del Valle de Teotihuacan. La población prehispánica*. Dirección de Talleres Gráficos, México D.F., edición de 1979 facsimilar edition.
- Goodall, C. (1991). Procrustes Methods in the Statistical Analysis of Shape. *Journal of the Royal Statistical Society. Series B (Methodological)*, 53(2):285–339.
- Gurt Esparraguera, J. M. and Buxeda i Garrigós, J. (1996). Metrologia, composició modular i proporcions de les basíliques cristianes de Llevant peninsular i de les Balears. In *Spania. Estudis d'Antiguitat Tardana oferts en homenatge al professor Pere de Palol i Salellas*, volume 12 of *Sèrie Il·lustrada, Biblioteca Abat Oliba*, pages 137–156. Publicacions de l'Abadia de Montserrat, Barcelona.
- Kendall, D. G. (1984). Shape Manifolds, Procrustean Metrics and Complex Projective Spaces. *Bulletin of the London Mathematical Society*, 16:81—121.
- Lawing, A. M. and Polly, P. D. (2010). Geometric morphometrics: recent applications to the study of evolution and development. *Journal of Zoology*, 280:1–7.

- Lele, S. R. (1993). Euclidean distance matrix analysis (EDMA) of landmarks data: estimation of mean form and mean form difference. *Mathematical Geology*, 25:573–602.
- Lele, S. R. and McCulloch, C. E. (2002). Invariance, Identifiability, and Morphometrics. *Journal of the American Statistical Association*, 97(459):796–806.
- Lele, S. R. and Richtsmeier, J. T. (2001). *An Invariant Approach to Statistical Analysis of Shapes*. Interdisciplinary Statistics. Chapman & Hall/CRC, Boca Raton.
- Lombardo de Ruiz, S. (1990). Conceptualización y abstracción: el arte teotihuacano. In *Arte pre-colombino de México*, pages 47–49. Olivetti-Electa, Milano.
- Lombardo de Ruiz, S. and Nalda, E. (1996). *Temas Mesoamericanos*. Colección Obra Diversa. INAH, México D.F.
- Magurran, A. E. (2004). *Measuring Biological Diversity*. Blackwell Publishing, Malden.
- MathSoft (1999). *S-PLUS 2000. User's Guide*. Data Analysis Products Division, MathSoft, Seattle.
- Michelet, D. and Allain, A. (2009). La sculpture à Teotihuacan. In *Teotihuacan. Cité des dieux*, pages 136–145. Musée du Quai Branly-Somogy Éditions d'Art, Paris.
- Millon, R. (1974). The study of urbanism at Teotihuacan, Mexico. In Hammond, N., editor, *Mesoamerican Archaeology: New Approaches*, pages 335–362. University of Texas Press, Austin.
- Morelos García, N. (2002). Las evidencias iconográficas del Complejo Calle de los Muertos. In Ruiz Gallut, M. E., editor, *Ideología y política a través de materiales, imágenes y símbolos. Memoria de la Primera Mesa Redonda de Teotihuacan*, pages 23–59. UNAM-CONACULTA-INAH, México.
- Mosimann, J. E. (1970). Size Allometry: Size and Shape Variables with Characterizations of Log-normal and Generalized Gamma Distributions. *Journal of the American Statistical Association*, 65(330):930–945.
- Nicholson, H. B. (2002). Major Sculpture in Pre-Hispanic Central Mexico. In Wauchope, R., editor, *Handbook of Middle American Indians. Archaeology of Northern Mesoamerica*, volume Part 1. X, pages 92–135. University of Texas Press, Austin.
- Oropeza, M. (1968). *Escultura teotihuacana*, volume 3 of *Colección breve*. Edimex, México D.F.
- Packard, G. C. and Boardman, T. J. (1999). The use of percentages and size-specific indices to normalize physiological data for variation in body size: wasted time, wasted effort. *Comparative Biochemistry and Physiology Part A*, 122:37–44.
- Pasztor, E. (1993a). An Image is Worth a Thousand Words: Teotihuacan and the Meanings of Style in Classic Mesoamerica. In Rice, D. S., editor, *Latin American Horizons*, Dumbarton Oaks Pre-Columbian Conference Proceedings, pages 113–146. Dumbarton Oaks Research Library and Collection, Washington, D.C.
- Pasztor, E. (1993b). The Natural World as Civic Metaphor at Teotihuacan. In Townsend, R. F., editor, *Ancient Americas: Art from Sacred Landscapes*, pages 135–146. The Art Institute of Chicago, Chicago.
- Pérez-Claros, J. A., Palmqvist, P., and Olóriz, F. (2002). First and Second Orders of Suture Complexity in Ammonites: A New Methodological Approach Using Fractal Analysis. *Mathematical Geology*, 34(3):323–343.

- Richtsmeier, J. T., DeLeon, V. B., and Lele, S. R. (2002). The Promise of Geometric Morphometrics. *Yearbook of Physical Anthropology*, 45:63–91.
- Richtsmeier, J. T., Lele, S. R., and Cole III, T. M. (2005). Landmark Morphometrics and the Analysis of Variation. In Hallgrimson, B. and Ball, B., editors, *Variation. A General Concept in Biology*, chapter 4, pages 49–69. Elsevier Academic Press, New York.
- Rohlf, F. J. (1990). Fitting curves to outlines. In Rohlf, F. J. and Bookstein, F. L., editors, *Proceedings of the Michigan morphometrics workshop*, number 2 in Special Publications, pages 167–177. University of Michigan Museum of Zoology, Ann Arbor.
- Sarro, P. J. (1988). *The monumental stone sculpture of Teotihuacan, Mexico: frontality and focus in urban spaces*. Master thesis, Columbia University, New York.
- Sarro, P. J. (1991). The Role of Architectural Sculpture in Ritual Space at Teotihuacan, Mexico. *Ancient Mesoamerica*, 2(2):249–262.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27:379—423 and 623—656.
- Slice, D. E. (2005). Modern Morphometrics. In Slice, D. E., editor, *Modern Morphometrics in Physical Anthropology*, Developments in Primatology: Progress and Prospects, chapter 1, pages 1–45. Kluwer Academic/Plenum Publishers, New York.
- Slice, D. E. (2007). Geometric Morphometrics. *Annual Review of Anthropology*, 36:261–281.
- Sugiyama, S. (2005). *Human Sacrifice, Militarism, and Rulership: Materialization of State Ideology at the Feathered Serpent Pyramid, Teotihuacan*. Cambridge University Press, Cambridge.
- Verrecchia, E. P. (2003). Foreword: Image Analysis and Morphometry of Geological Objects. *Mathematical Geology*, 35(7):759–762.
- Villalonga Gordaliza, A. (2010). Un acercamiento a la litoescultura en Teotihuacan, México. Antiguas problemáticas, nuevas propuestas. In *Jornadas de Jóvenes en Investigación Arqueológica (5-7 mayo de 2010)*. Universitat Autònoma de Barcelona, Bellaterra. In press.
- Villalonga Gordaliza, A. (2011). La escultura antropomorfa en piedra de Teotihuacan, México: problemáticas, antecedentes y nuevas propuestas. In Dalla Corte, G. and García Jordán, P., editors, *Sociedades diversas, Sociedades en cambio. América latina en perspectiva histórica. XII Encuentro-Debate América Latina ayer y hoy*, pages 112–121. Sección Departamental de Historia de América y África de la Universidad de Barcelona, Barcelona.