Geometric morphometrics are a class of multivariate methods for measuring and analyzing the shapes of objects. Geometric morphometrics are distinctive in that they use Cartesian coordinates \((x, y, z)\) of landmark points in place of linear measurements as variables (Bookstein 1991). Most geometric morphometric methods use Procrustes superimposition to remove scale, rotation, and translation from the coordinates, which eliminates confounding effects of size, but which limits analyses to rigid objects. A wide range of multivariate statistical tests can be employed to assess the relationship of shape to other variables of interest (Dryden and Mardia 1998). One of the most popular features of geometric morphometric methods are that analytical results can be presented visually as deformations of objects instead of as tables of numbers.

Methodology

The hallmark of geometric morphometrics is that points on the surfaces of objects (or in their interiors, if volumetric data such as CT or MRI scans are available) are used to represent their shape. For example, the shape of a square might be represented by points at its four corners, or a Paleolithic hand axe might be represented by a series of points traced around its perimeter (see Figure 1a). The points may be two-dimensional (2D) or three-dimensional (3D), producing two \((x, y)\) or three \((x, y, z)\) variables respectively per point. In any one analysis, the same number of points must be placed on every object in corresponding locations and in the same order.

Three strategies for placing points can be distinguished. Landmarks, in the strict sense, are single points, each of which represents the position of a particular structure. For biological objects, landmarks are usually placed at biologically homologous positions such as the junction between three bones or the openings of foramen (see Figure 1a, large circles). Landmarks can also be placed geometrically, for example at the point of maximal depth of the mandibular ramus. Semilandmarks are series of points arranged geometrically on a structure. Edges or curves can be represented by a string of semilandmarks, traced like a dotted line around them (see Figure 1a, small circles), and surfaces can be covered with a three-dimensional grid or cloud of semilandmarks. Like with landmarks, the ordering of semilandmarks on each object must correspond. For example, the 100 semilandmarks on the hand axes in Figure 1 all start at the midline of the base and trace the perimeter clockwise. Semilandmarks can be placed at regular intervals, or their placements can be optimized by sliding them until the shape differences between objects is minimized, a strategy known as "sliding semilandmarks" (Gunz, Mitteroecker, and Bookstein 2005).

Regardless of which point-placing strategy is used, point coordinates must be registered in a common coordinate system before analysis can proceed. Most geometric morphometric methods use generalized Procrustes superimposition to register coordinates (Rohlf and Slice 1990). In Procrustes superimposition, the coordinates of all objects are scaled, translated, and rotated to minimize the shape differences between them. Each object is centered on the origin \((x = 0, y = 0, z = 0)\), scaled so that the squared distances between each point and the origin sum to 1 (which is known as unit centroid size), and the objects are rotated about the origin until the Procrustes distance, or sum of squared distances between corresponding points between two objects, is minimized. This least-squares-fitting algorithm ensures that differences in shape are minimized or, in other words, that they are not due to spurious differences in orientation within the coordinate system, which is an important prerequisite for applying statistical tests for differences in shape. Procrustes superimposition has the statistical consequence of reducing the
degrees of freedom in the data by removing the effects of scale, translation, and rotation. Two-dimensional data are left with $2k-4$ degrees of freedom, where $k$ is the number of landmark points, and three-dimensional data are left with $3k-7$ degrees of freedom. If the number of objects ($n$) in the analysis is less than these expectations, then the number of degrees of freedom is $n-1$.

Once superimposed, Procrustes coordinates are usually transformed into shape variables for further statistical analysis. Ideal shape variables are equal in number to the degrees of freedom and are orthogonal to (uncorrelated with) one another, but preserve shape differences between objects. The most common kind of shape variables are principal components (PC) scores (see Figure 1b). When based on the covariance method, principal components ordination represents a rigid geometrical transformation of objects to a new set of orthogonal axes whose number equals the statistical degrees of freedom of the data set. The coordinates of objects in the PC space, which are known as scores, are the shape variables. For example, the scores of Hand Axe 1 in Figure 1b are $-0.08$ on PC 1 and $0.00$ on PC 2. Another commonly used shape variable are the partial warps scores, which are derived from an ordination of specimens based on the so-called bending energy matrix derived from spatial differences in the locations of points on the objects.

The principal components space is often viewed as a morphospace. Each point in the multidimensional morphospace corresponds to a particular configuration of landmarks. Objects are placed at a given point when their shape corresponds to its particular configuration. Distances between objects in the morphospace are equal to the Procrustes distances between their superimposed landmark points if measured as a Euclidean distance calculated from the full set of PC scores. Principal component shape spaces are sample dependent in that their coordinate system is centered on the mean shape (the shape configuration represented by the point at the origin of the PC axes is the arithmetic mean of the Procrustes superimposed coordinates) and in that the first PC is oriented along the major axis of variation in the sample. If new objects are added, the distance between objects does not change, but the ordination of the axes and the values of the scores do.

Statistical tests and other kinds of analysis are performed on the matrix of shape variables. For the most part, any kind of multivariate statistical method can be applied to shape data, so long as the implications of the loss of degrees of freedom are taken into account (Dryden and Mardia 1998). Multivariate regression can be used to assess the relationship between shape and a continuous variable. For example, age-dependent changes in skull shape would be assessed by regressing shape scores derived from landmarks of the skull onto the chronological age of the individuals in the analysis. Multivariate analysis of variance (MANOVA) can be used to test for differences

Figure 1  (a) Paleolithic hand axe with 60 semilandmarks placed at regular intervals around the perimeter (small circles) and two landmarks at the midline of the base and point (large circles). (b) The first two dimensions of a principle components analysis of the Procrustes superimposed coordinates of three hand axes. Axe 1 is the tool shown in (a). (c) A thin-plate spline diagram showing the shape difference between Axe 2 and Axe 1 as a deformation of the coordinates of the former into the latter. Modified from Smith, 1912.
between groups. For example, MANOVA can be used to test hypotheses that morphology is sexually dimorphic or that a new fossil sample is distinct from previously described species. Because shape variation often fails to meet the assumptions of traditional parametric statistics and because sample sizes in anthropology are often necessarily unbalanced, non-parametric permutation and Monte Carlo tests are to be preferred for shape analysis. When the objects being analyzed belong to more than one biological population or species, phylogenetic comparative statistical methods should be considered (Harvey and Pagel 1991).

Because the values of PC shape scores are sample dependent, it is perilous to try to interpret the causal relationship between just one principal component and an external factor (Polly and Motz 2016). For example, the arrangement of species along PC 1 in a shape analysis of tooth shape in monkeys may appear to correspond to the proportion of vegetation in their diets, but it is unlikely that the true relationship between tooth shape and diet corresponds precisely to that axis. Instead, a fully multivariate regression technique that utilizes the full set of shape scores from all PCs should be employed, not a visual inspection (or even a statistical analysis) of just one or two PCs.

Visualization techniques

Because shape variables directly correspond to Cartesian coordinates of points on objects, results from geometric morphometric analysis can often be presented as graphical depictions of the objects themselves. The difference between the shape of two objects, for example, can be illustrated as the displacement of landmark points. Interpolation methods such as thin-plate splines can use the landmark displacements to deform a grid in order to represent the deformation of one shape into another, a technique that has become one of the hallmarks of geometric morphometric methods (Bookstein 1991). For example, Figure 1c shows the difference between hand axes 1 and 2 as a thin-plate spline grid deformation of axe 2 into axe 1, which requires compression around the distal end to produce the latter's tapered outline. While this example is applied to two objects, the same strategy can be used to show the difference between two group means (e.g., shape differences between sexes). Thin-plate spline grids are most effective with two-dimensional data. Results for three-dimensional data are often shown as animations of the change in shape along a linear trajectory through morphospace. Results from multivariate regression can be depicted the same way since the slope and intercept parameters it produces describe a line through morphospace.

History of geometric morphometrics

Geometric morphometric methods in the strict sense were first described in the late 1980s and early 1990s. Multivariate ordination and statistical methods had already been common since the 1960s (e.g., Blackith and Reyment 1971) and were applied to matrices of data such as linear measurements, angles, and areas. Statistical analysis of relationships between shape and other variables was possible using these methods, but interpretation could be difficult because size and shape were comingled and because the mapping of results onto biological form was usually interpreted by visual inspection of loadings of variables onto statistical factors. In a conscious attempt to capture the graphical power of the qualitatively estimated deformation grids that D’Arcy Thompson popularized in his book On Growth and Form (Thompson 1917), Fred Bookstein and others began experimenting with truss analysis of linear measurements that systematically connected landmarks points in order to be able to transform factor loadings into a graphical map of which parts of the shape contributed most to differences between objects. The breakthrough came with the realization that Cartesian coordinates could be substituted for linear measurements and that thin-plate spline interpolations could then be used to produce deformable grids (Bookstein 1991). The addition of Procrustes superimposition as a tool to minimize shape differences as a prerequisite for statistical analysis (Rohlf and Slice 1990) completed the “geometric morphometric revolution.”
GEOMETRIC MORPHOMETRICS

Shortcomings

While geometric morphometric methods have many advantages over measurement-based methods, including the ability to mathematically factor out size and the ease with which results can be depicted graphically, they also have stubborn limitations.

Arguably the most constraining limitation is that the methods can only be applied to rigid objects. For the relative positions of Cartesian coordinates to be meaningful, each point used in an analysis must occupy a fixed position relative to all other points. The consequence is that objects that have mobile relationships to one another, such as mandible and cranium or body and limbs, cannot easily be analyzed with geometric morphometrics because spurious differences in their relative positions contribute to shape differences (e.g., the same individual would have a different “shape” if the limbs were moved to new positions). Special techniques such as two-block partial least squares can be used to analyze covariation in shape between objects that are not fixed, such as mandible and cranium, but these become cumbersome for systems involving more than two elements.

Another limitation is that the displacements of landmark points estimated by geometric morphometric methods are relative, not absolute, because of Procrustes superimposition. Consequently, it is impossible to localize which parts of a shape have changed. This issue has been exemplified by a thought problem involving the fairy-tale puppet Pinocchio, whose nose grew every time he told a lie. The only change in the shape of Pinocchio’s profile would be elongation of his nose; however, Procrustes superimposition distributes the differences across all landmarks because of the centering and minimization components of its algorithm. Geometric morphometrics can accurately recover the fact that the relationship between the nose and the rest of the head changed, but it would not be able to distinguish between growth of the nose, shrinking of the head, or an intermediate combination of the two. Alternative methods like Euclidean distance matrix analysis (EDMA), which is similar to the truss analysis described above, have been developed specifically to localize which parts of a shape are changing.

Applications in anthropology and archaeology

Geometric morphometrics have been applied to many problems in anthropology and archaeology. The techniques are heavily used in evolutionary paleoanthropology, where they have been used to test for shape differences between putative hominin species, to statistically test hypotheses about functional relationships between morphology and factors like diet or sex, and to reconstruct ancestral morphologies using tip taxa and phylogenetic trees. Geometric morphometrics have also been applied to human-produced objects such as stone tools and floor-plan configurations of structures such as fortification towers and tipi rings, as well as to objects as varied as geographic topography and footprint outlines.

Future directions

The ease with which three-dimensional data can be obtained with laser scanners, computed tomography (CT) scanners, magnetic resonance imaging (MRI), and Lidar, have increased the demand for geometric morphometric analyses of fully three-dimensional surfaces. At the same time, they have exposed the shortcomings imposed by the requirement that all objects in a geometric morphometric analysis must have the same number of points in the same topographical configuration. Only a few algorithms exist for placing semilandmarks on three-dimensional surfaces and all of them are expensive in terms of the work and computational power required to implement them. The greatest advances in geometric morphometrics over the next decade are likely to be in automated algorithms that match surfaces and apply point grids. Some recent advances even suggest the demise of geometric morphometrics as we know it, in favor of strategies that map surfaces onto one another using pure mathematical functions instead of landmark points.
SEE ALSO: Analysis of Variance and Covariance; Cluster Analysis; Cranio-ometry; Dental Metrics; Human Skeletal Data Quantification; Inferential Statistics; Three-Dimensional (3D) Facial Reconstruction; Lithics Data Quantification

REFERENCES


FURTHER READINGS