

# Comparing shapes, understanding evolution

Guillermo Sapiro<sup>1</sup>

Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455

The study of shape has intrigued some of the brightest minds of humanity, from Leonardo Da Vinci and Carl Friedrich Gauss to some of the top scientists of the modern era. The mathematics to analyze shapes are both beautiful and challenging, covering a variety of tools, from topology (1), to metric and differential geometry (2, 3), to statistics (4). The applications are very diverse and potentially life-changing, and they range from brain research to structural biology and archeology. At the core of this research area is shape comparison: determining how to match, compare, and compute the distance between pairs of shapes. This is the topic addressed by the work of Boyer et al. (5) in PNAS, with a unique interdisciplinary team of anthropologists, archaeologists, computer scientists, and mathematicians.

Understanding why comparing shapes is so important is better illustrated with a number of examples. Let us start with brain research, where, for example, we are interested in investigating how the brain changes as we grow (6) or how a normal brain compares with that of a patient who has Alzheimer's disease. In this case, the shapes can be obtained from MRI and can, for example, represent the gray-white matter boundary, thereby challenging us to find maps and distances between such 2D surfaces (7). In HIV research in structural biology, data can be obtained via cryotomography, and it is important to understand the underlying shape of the envelope glycoproteins that mediate virus binding to initiate infection and how this shape changes, for example, in the presence of antibodies, information critical for the development of a vaccine. Shape comparison and matching are fundamental to compute the conformations of such protein complexes (8). Boyer et al. (5) present unique applications in anthropology, archeology, and evolution, as discussed below. Let us not forget that our surrounding world is composed of shapes and that shape analysis is critical to navigate it as well as to develop automatic systems capable of emulating basic human performance, such as answering the simple question "Is this a chair?".

Comparing shapes is difficult because of the intrinsic complexity of shapes in nature (e.g., proteins, human brains) as well as the large variability encountered within shape classes. Although simple characteristics to compare shapes, such as volume, can already provide valuable information,



**Fig. 1.** Finding corresponding/matching points between shapes is a very challenging problem and is critical for a number of applications (A, with kind permission from Springer Science+Business Media: *International Journal of Computer Vision*, A Gromov-Hausdorff framework with diffusion geometry for topologically-robust non-rigid shape matching, Vol 89, 2009, pp 266–286, A. M. Bronstein, M. M. Bronstein, M. Mahmoudi, R. Kimmel, and G. Sapiro, Fig. 11; B, manufactured by Plan Toys, Thailand).

more sophisticated features and distances are needed most often. These distances can be derived following the computation of a correspondence between the shapes: a map between points in the shapes being compared. Computing such correspondence, and, from it, distances, is the essence of the work of Boyer et al. (5) and of much of the literature in this area in recent years (finding correspondence has applications beyond shape comparison and is critical, for example, in morphing, as exemplified by the famous Michael Jackson musical video *Black or White*).

Finding an appropriate correspondence between shapes is often addressed by considering a discrete set of landmarks or corresponding points or curves. This approach is common, for example, in Procrustes analysis, a form of statistical shape analysis that derives its name from the mythological Greek rogue who made his victims fit his bed by stretching their limbs or cutting them off, and has been very popular in brain matching (7). Some landmarks might be natural for some classes of shapes (e.g., tip of the nose for faces) but are not for others. Even if they are easy to define and universally acceptable by the corresponding community (which is often not the case), marking them requires having experts in the field or developing advanced computational techniques (which are often problematic by themselves); is very time-consuming, and thereby forbidden for large datasets; is subject to much subjectivity in their selection; and is also prone to errors and contamination. Avoiding landmarks altogether is thereby desirable, as done by Boyer et al. (5) based on a combination of beautiful and computable mathematical structures, including Monge–Kantorovitch mass transportation theory and conformal maps. The Monge–Kantorovitch theory follows from work of the French mathematician Gaspard Monge 3 centuries ago and the Soviet mathematician Leonid Kantorovitch in the past century, and it relates to the study of optimal transportation and resources allocation problems. Conformal maps are angle-preserving and are familiar to the readers because they are often used in cartography to map the round earth onto the plane.

Author contributions: G.S. wrote the paper.

The author declares no conflict of interest.

See companion article on page 18221.

<sup>1</sup>E-mail: guille@umn.edu.

It is interesting to see how old theories and practices come back to life in modern mathematics and applications.

Boyer et al. (5) derive, via energy/variational formulations, an extension of the conformal Wasserstein distance, which is the mass transport problem for the conformal factors (area distortions produced by the conformal maps) corresponding to two given surfaces. This modification, based on performing local computations, is necessary, in part, because the original classic formulation is computationally very complex. This newly developed computationally feasible neighborhood distance has interesting properties demonstrated in complementary theoretical supporting information provided by the authors.

The proposed neighborhood conformal Wasserstein distance is isometric-invariant, meaning it is intrinsic and invariant to bends and it preserves distances on the surface, suitable for deformations, such as those in Fig. 1A. A different distance, the continuous Procrustes distance, is investigated, which is also landmark-free and considers extrinsic properties more suitable for rigid matches, such as those in Fig. 1B. The interplay between intrinsic and extrinsic properties of shapes is an active area of research, and it has often been demonstrated that both classes of maps are important.

The beautiful mathematical entities developed and studied by Boyer et al. (5) would have limited value if they were not efficiently computable. To complete the technical contribution of the work, the

authors study such issues. In particular, they show that by limiting the space over which the optimization is performed, very efficient algorithms can be developed. In the companion theoretical paper and supporting information, they carefully analyze the bounded differences obtained by

## Boyer et al. present unique applications in anthropology, archeology, and evolution.

such optimization-motivated restriction, showing that it does not hurt; computational details are also provided in the supporting information.

Boyer et al. (5) apply these mathematical and computational developments in three different teeth and bone datasets of human and nonhuman skeletal anatomy, presenting a real application in comparative and evolutionary morphology. Digitized 2D surfaces embedded in three dimensions were obtained from high-resolution X-ray computed tomography scans of 116 second mandibular molars of prosimian primates and nonprimate close relatives; 57 proximal first metatarsals of prosimian primates, New and Old World monkeys; and 45 distal radii of apes and humans. For every pair of surfaces, the proposed algorithms provide the correspondence map and a nonnegative number giving their distance or dissimilarity. The results

of these automatic algorithms are compared with standard tools used by morphologists. The comparisons are done at both the quantitative level (Mantel significance) and qualitative level (observing symmetry in combined distance matrices).

In addition, Boyer et al. (5) perform a taxonomic classification task and show that their automatic continuous Procrustes distance performs as well as the standard manual (landmark-based) one. Finally, because their automatic matching technique provides the actual map and not just the distance, they compare the manual landmarks with those obtained by their technique, showing once again the value and accuracy of their automatic framework. Beyond this, the automatic and objective propagation of landmarks via the computed maps can shed some light on comparative morphology and the evolutionary relatedness among groups of organisms, as illustrated by Boyer et al. (5) with a study of different alternative evolution paths among different taxonomic groups. The study concentrates on mapping landmarks from a *Microcebus* tooth onto a *Lepilemur* tooth, showing possible direct and indirect evolutionary paths.

The work by Boyer et al. (5) has the added value of serving as an example of interdisciplinary research, with contributions and motivations from multiple disciplines. As such, it opens the door to applications of the proposed shape distances and matching techniques in other areas of science, as well as to additional mathematical questions coming from evolutionary morphology.

1. Edelsbrunner H (2009) *Computational Topology* (American Mathematical Society, Providence, RI).
2. Gromov M (2006) *Metric Structures for Riemannian and Non-Riemannian Spaces* (Birkhäuser, Boston).
3. Spivak M (1999) *A Comprehensive Introduction to Differential Geometry I-V* (Publish or Perish, Houston, TX).
4. Kendall DG (1989) A survey of the statistical theory of shape. *Stat Sci* 4:87–120.
5. Boyer DM, et al. (2011) Algorithms to automatically quantify the geometric similarity of anatomical surfaces. *Proc Natl Acad Sci USA* 108:18221–18226.
6. Thompson PM, et al. (2000) Growth patterns in the developing human brain detected by using continuum mechanical tensor maps. *Nature* 404:190–193.
7. Toga A, ed (1998) *Brain Warping* (Academic Press, New York).
8. Bartesaghi A, Subramaniam S (2009) Membrane protein structure determination using cryo-electron tomography and 3D image averaging. *Curr Opin Struct Biol* 19: 402–407.