COMMENTARY

From understanding of color perception to dynamical systems by manifold learning

Ron Kimmel[a,1]

Comprehending Color

Let us start with a seemingly unrelated field to that described in the article by Yair et al. (1) in PNAS. The field of psychophysics deals with the relationships between physical stimuli and mental phenomena. An excellent example is the scientific community's early efforts to study the human perception of color. Scientists have been intrigued by visual awareness of colors, trying to understand our interpretation of colors and attempting to quantify human perception with simple equations. Roughly speaking, one could divide these efforts into axiomatic ones that gave birth to the Young, Maxwell, Helmholtz, and, later on, Schrödinger so-called "inductive color line elements" and the empirical color arc-lengths that reflected the effort to virtually embed measurements of human color perception into a simple, often Euclidean, domain. In fact, the latter school of thought, of treating the problem empirically rather than axiomatically, is probably one of the earliest attempts to apply a manifold learning technique to study a psychophysical phenomenon. The outcome was the insightful observation that human color perception is 3D, while most birds probably have (and most dinosaurs probably had) a color perception manifold of higher dimensions and most other mammals share a lower dimensional space for the (lack of) perception of color. One of the analysis tools used to arrive at this important observation is known as multidimensional scaling (MDS), and is related to the famous principal component analysis machinery that is commonly used in big data representation, for which various modern generalizations exist. While axiomatic realizations of studying the color receptors in the eye lead Maxwell (2) to the understanding that color images could be synthesized by a linear combination of three monochromatic colors, the space in which one should operate and the ways by which anchor (basis) colors should be selected has been the topic of many scientific and industrial explorations leading to the design of modern mobile, computer, and television screens, as well as almost all printing devices. The early empirical analysis of color perception by manifold learning is indeed a remarkable step in our ability to model human behavior and harness this understanding to our benefit.

When trying to process images so as to enhance them and improve their quality, the color line element should obviously come into play. Geometry modeling of image formation indeed led researchers to the introduction of a new manifold that marries the color line element with the image coordinates, giving rise to a 2D manifold (the image) embedded in a 5D space, where three of these dimensions are an exact result of our understanding of color perception. Indeed, these geometric observations are incorporated, in one way or another, into most modern color image processing tools. The Beltrami filter, bilateral filter, and, in fact, most color processing methodologies exploit our understanding of the fundamental geometry behind color perception in one way or another.

Image Understanding

Once our perception of color was well understood, automatic identifying and classifying of the content in an image became the next problem the academic community was trying to resolve. This effort gave birth to a field known as robot vision, computer vision, or image understanding. The notion of so-called "deep learning" without "understanding" has very recently substantially changed the way in which fields like computer vision are addressed. Within the context of deep learning, we can now deliberately avoid the term understanding, as deep learning methods try to interpolate between given observations, fitting a nonlinear interpolation mechanism with a multiparameter fixed architecture into a huge corpus of annotated observations. In these research arenas, "big data" of annotated objects are usually publicly available, which make the process of tuning the large number of parameters of the interpolation mechanism (often referred to as a neural network) feasible. At the other end, representation of empirical observations of physical phenomena for which there is no apparent explanation through a nonlinear multiparameter generic mechanism without an Occam's razor simple theoretical or axiomatic reasoning is what the article by Yair et al. (1) is about.

[a]Computer Science Department, Technion-Israel Institute of Technology, Haifa 32000, Israel

Author contributions: R.K. wrote the paper.

The author declares no conflict of interest.

See companion article on page E7865.

[1]Email: ron@cs.technion.ac.il.

A couple of centuries after the manifold learning way of deciphering the psychophysical modeling of color perception, the model of Yair et al. (1) comes in a timely fashion, introducing novel concepts and ideas by which some understanding could still be extracted from the learned objective. In this case, the learning is not deep by its modern definition but rather belongs to the class of approaches known as manifold learning, or geometric reasoning.

Study of Shapes and Forms

Another interesting and related geometric reasoning mechanism for analyzing the structure of the cortical surface by mapping its geometric structure into a Euclidean one was introduced by Schwartz et al. (3). Treating surfaces as metric spaces was an interesting new line of thought that came about from the effort of Schwartz et al. (3) to find a unified parametrization for the surface of the brain. The idea was to first embed the geometric structure of the surface into a 2D Euclidean space invoking, yet again, the MDS procedure. This effort, in a sense, is a generalization of the celebrated mapmakers' cartography problem.

In an attempt to refine the shape-matching problem, the Gromov–Hausdorff distance, which was thought to be restricted to theoretical analysis of abstract metric spaces, has also been adopted by the shape analysis and manifold learning communities. It was suggested as a candidate for measuring the discrepancy between two deformable shapes (4), followed by the generalized MDS (GMDS) (5) that numerically approximates that measure to compute the map between two surfaces that best preserves corresponding intergeodesic distances. In the GMDS framework, the Hausdorff distance was replaced by the Wasserstein norm. Other efforts that try to find a universal parametrization for surfaces suggested conformal mapping to a disk (6–8) and even using the earth mover's distance (EMD) between corresponding conformal factors on the disk (9) in a fashion related to that suggested in the model of Yair et al. (1).

Distances measured on surfaces are smooth functions for which the gradient magnitude is equal to 1 almost everywhere. As such, they are suited for compact spectral representation (10), allowing us to translate the Gromov–Hausdorff framework into the spectral domain. In fact, resorting to a truncated spectral domain as an alternative for computing the Gromov distance explicitly for metric analysis was probably first suggested by Bérard et al (11) and more recently adopted to surface analysis by Ovsjanikov et al. (12).

Deciphering Dynamical Systems

In their paper, Yair et al. (1) are applying geometric manifold learning techniques to find compact representations of empirical observations of physical phenomena through a geometry known as diffusion maps. That is, distances between events are described by kernels defined through some basis functions. On top of that, Wasserstein distance is efficiently approximated using recent developments with an even smoother version of measuring discrepancies, treating distances between events as ones between probability distributions that relate to the Monge–Kantorovich problem or the EMD. In the paper of Yair et al. (1), the EMD is computed in such a way that the coupling between the "ground distance" is constructed in a data-driven way from the observations. Although most of the ingredients and tools have been known for a while, their combination allows the authors to nicely apply them to find the governing behavior of a pendulum extracted from a video of its different positions in time as one example and to find a compact geometric description of even more complicated systems, like two pendula. Capturing the behavior of dynamical systems efficiently for which only the time axis is a common denominator poses a challenge that Yair et al. (1) seem to gracefully overcome. It is yet to be seen if this novel application of manifold learning to dynamical systems would lead to technological innovative inventions such as those resulting from deciphering human perception of color.

A couple of centuries after the manifold learning way of deciphering the psychophysical modeling of color perception, the model of Yair et al. comes in a timely fashion, introducing novel concepts and ideas by which some understanding could still be extracted from the learned objective.

1 Yair O, Talmon R, Coifman RR, Kevrekidis IG (2017) Reconstruction of normal forms by learning informed observation geometries from data. *Proc Natl Acad Sci USA* 114:E7865–E7874.
2 Maxwell JC (2003) *The Scientific Papers of James Clerk Maxwell*, ed Niven WD (Dover Publications, Mineola, NY), Vol 1.
3 Schwartz EL, Shaw A, Wolfson E (1989) A numerical solution to the generalized mapmaker's problem: Flattening nonconvex polyhedral surfaces. *IEEE Trans Pattern Anal Mach Intell* 11:1005–1008.
4 Memoli F, Sapiro G (2005) A theoretical and computational framework for isometry invariant recognition of point cloud data. *Found Comput Math* 5:313–347.
5 Bronstein AM, Bronstein MM, Kimmel R (2006) Generalized multidimensional scaling: A framework for isometry-invariant partial surface matching. *Proc Natl Acad Sci USA* 103:1168–1172.
6 Gu X, Wang Y, Chan TF, Thompson PM, Yau ST (2004) Genus zero surface conformal mapping and its application to brain surface mapping. *IEEE Trans Med Imaging* 23:949–958.
7 Lipman Y, Funkhouser T (2009) Möbius voting for surface correspondence. *ACM Trans Graph* 28:72.
8 Zeng Y, et al. (2010) Dense non-rigid surface registration using high-order graph matching. *IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, San Francisco), pp 382–389.
9 Lipman Y, Daubechies I (2011) Conformal Wasserstein distances: Comparing surfaces in polynomial time. *Adv Math* 227:1047–1077.
10 Aflalo Y, Kimmel R (2013) Spectral multidimensional scaling. *Proc Natl Acad Sci USA* 110:18052–18057.
11 Bérard P, Besson G, Gallot S (1994) Embedding Riemannian manifolds by their heat kernel. *Geom Funct Anal* 4:373–398.
12 Ovsjanikov M, Ben-Chen M, Solomon J, Butscher A, Guibas L (2012) Functional maps: A flexible representation of maps between shapes. *ACM Transaction on Graph* 31:30:1–30:11.