

The Manifold Ways of Perception

H. Sebastian Seung and Daniel D. Lee

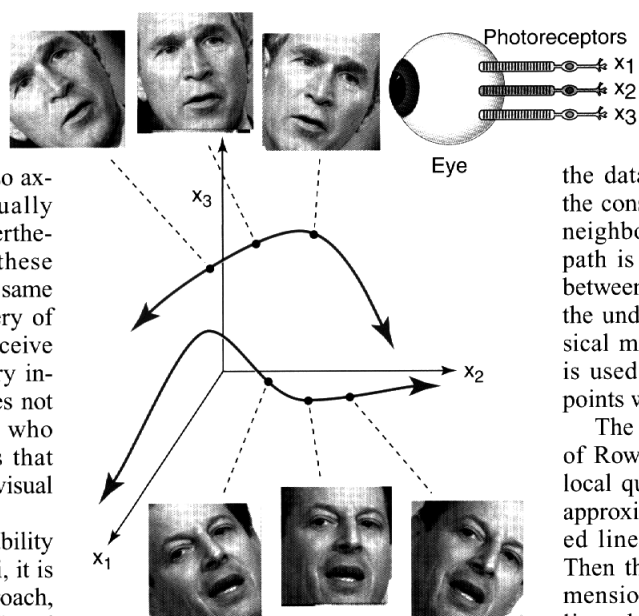
Two-and-a-half millennia ago, the Greek philosopher Heraclitus, observing that the world is in eternal flux, wrote that you can never step in the same river twice. If he were alive today and working as a psychologist, he might say that you can never see the same face twice. Indeed, faces can grow hair, acquire wrinkles, or be surgically enhanced. But facial images also vary from moment to moment, as you can demonstrate at home while watching television. Make a small aperture in a piece of paper, and place it over a face on the screen. The light coming through the aperture will vary with time, mostly as a result of changes in the location and orientation of the face.

The aperture might show a tooth at one instant, and a nostril at the next, crudely simulating the fluctuations in light incident on a single retinal photoreceptor cell. This illustrates that the signals carried from the eye to the brain by the million or so axons in the optic nerve are perpetually changing as we look at a face. Nevertheless, we are able to perceive that these changing signals are produced by the same object. This is the fundamental mystery of perception: How does the brain perceive constancy even though its raw sensory inputs are in flux? The mystery intrigues not only scientists but also engineers, who yearn to construct vision machines that equal the performance of humans at visual object recognition.

To precisely characterize the variability of images and other perceptual stimuli, it is essential to take a mathematical approach, which is just what Tenenbaum *et al.* (1) and Roweis and Saul (2) have done on pages 2319 and 2323 of this issue, respectively. An image can be regarded as a collection of numbers, each specifying light intensity at an image pixel. But a collection of numbers also specifies the Cartesian coordinates of a point with respect to a set of axes. Therefore, any image can be identified with a point in an abstract *image space*.

Now consider a simple example of image variability, the set M of all facial images generated by varying the orientation of a face (see the figure). This set is a continuous

curve in the image space. It is continuous because the image varies smoothly as the face is rotated. It is a curve because it is generated by varying a single degree of freedom, the angle of rotation. In other words, M is intrinsically one-dimensional, although it is embedded in image space, which has a high dimensionality equal to the number of image pixels. If we were to allow other types of image transformations, such as scaling and translation, then the dimensionality of M would increase, but would still remain far less than that of the image space. In this generalized case, M is said to be a *manifold* embedded in the image



Manifolds in visual perception. The retinal image is a collection of signals from photoreceptor cells. If these numbers are taken to be coordinates in an abstract image space, then an image is represented by a point. Only three dimensions of the image space are depicted, but actually the dimensionality is equal to the number of photoreceptor cells. As the faces are rotated, they trace out nonlinear curves embedded in image space. If changes in scale, illumination, and other sources of continuous variability are also included, then the images would lie on low-dimensional manifolds, rather than the simple one-dimensional curves shown. To recognize faces, the brain must equate all images from the same manifold, but distinguish between images from different manifolds. How the brain represents image manifolds is as yet unknown. According to one hypothesis, they are stored in the brain as manifolds of stable neural-activity patterns.

space. A curve is an example of a one-dimensional manifold, whereas a sphere is an example of a two-dimensional manifold (3).

Although the preceding discussion is biased toward vision, manifolds are also relevant to other types of perception. Furthermore, scientists in many fields face the problem of simplifying high-dimensional data by finding low-dimensional structure in it. Therefore, the manifold learning algorithms described by Tenenbaum *et al.* (1) and Roweis and Saul (2) are of potentially broad interest. The goal of the algorithms is to map a given set of high-dimensional data points into a surrogate low-dimensional space. Both start with a preprocessing step that decides for each data point which of the other data points should be considered its neighbors. Then both compute measures of the local geometry of the manifold, after which the original data points are no longer needed.

In the Isomap algorithm of Tenenbaum *et al.*, the local quantities computed are the distances between neighboring data points. For each pair of non-neighboring data points, Isomap finds the shortest path through the data set connecting them, subject to the constraint that the path must hop from neighbor to neighbor. The length of this path is an approximation to the distance between its end points, as measured within the underlying manifold. Finally, the classical method of multidimensional scaling is used to find a set of low-dimensional points with similar pairwise distances.

The locally linear embedding algorithm of Roweis and Saul computes a different local quantity, the coefficients of the best approximation to a data point by a weighted linear combination of its neighbors. Then the algorithm finds a set of low-dimensional points, each of which can be linearly approximated by its neighbors with the same coefficients that were determined from the high-dimensional data points. Both algorithms yield impressive results on some benchmark artificial data sets, as well as on "real world" data sets. Importantly, they succeed in learning nonlinear manifolds, in contrast to algorithms such as principal component analysis, which can only learn linear manifolds.

Because manifolds are fundamental to perception, the brain must have some way of representing them. Clues to the nature of this representation may come from studies of how information is encoded in large populations of neurons. Population activity is typically described by a collection of neural firing rates, and so can be represented by a point in an abstract space with dimensionality equal to the number

H. S. Seung is at the Howard Hughes Medical Institute and Brain and Cognitive Sciences Department, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. D. D. Lee is at Bell Labs, Lucent Technologies, Murray Hill, NJ 07974, USA.

of neurons. Neurophysiologists have often found that the firing rate of each neuron in a population can be written as a smooth function of a small number of variables, such as the angular position of the eye (4) or direction of the head (5). This implies that the population activity is constrained to lie on a low-dimensional manifold.

What is the connection between such neural manifolds and the image manifolds we have just discussed? According to a well-known idea, memories are stored in brain dynamics as stable states, or dynamical

attractors (6). Because the possible images of an object lie on a manifold, it has been hypothesized that a visual memory is stored as a manifold of stable states, or a continuous attractor (7). Recent studies of neural manifolds suggest that continuous attractors actually do exist in the brain (8, 9). Whether they are the basis of visual and other types of perception remains to be resolved. If the answer is affirmative, then manifolds will prove to be crucial for understanding how perception arises from the dynamics of neural networks in the brain.

References

1. J. Tenenbaum, V. de Silva, J. C. Langford, *Science* **290**, 2319 (2000).
2. S. Roweis, L. Saul, *Science* **290**, 2323 (2000).
3. K. Devlin, *Mathematics: The Science of Patterns* (Scientific American Library, New York, 1997).
4. J. L. McFarland, A. F. Fuchs, *J. Neurophysiol.* **68**, 319 (1992).
5. J. S. Taube, *Prog. Neurobiol.* **55**, 225 (1998).
6. J. J. Hopfield, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2554 (1982).
7. H. S. Seung, *Adv. Neural Info. Proc. Syst.* **10**, 654, (1998).
8. H. S. Seung, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 13339, (1996).
9. K. Zhang, *J. Neurosci.* **16**, 2112 (1996).