PATTERNS AND FEATURES AS MANIFOLDS: FACES IN IDENTIFICATION AND CATEGORIZATION

James F. Townsend & Michael J. Wenger
Indiana University

Abstract

The present chapter considers some of the possibilities associated with the use of infinite dimensional spaces in perceptual and cognitive theorizing. Of particular import and application are questions in face perception and cognition, a domain in which the use of infinite dimensional spaces would appear particularly advantageous, given the complexities (geometric, perceptual, and cognitive) associated with faces (more generally, heads). We outline suggestions for initial applications, both mathematical and experimental.

Despite the numerous examples in perception and cognition in which the objects of information processing could not be simple finite dimensional vector spaces, much of our theory and technology of measurement, scaling and psychophysics is devoted almost entirely to these. (Interestingly, as an aside, outside of the assumption of orthogonality and imposition of a metric, the vital assumptions of vector spaces are hardly ever employed; but that is beyond our present scope of discussion.) Townsend and Thomas (1993) provide references to exceptions to this general rule. In no way do we disparage the great importance of these spaces. However, we do believe it is high time to enlarge our perspective to include some of the rich possibilities that modern mathematics offers.

As an example that is intended to guide ourselves as well as to illustrate what we mean, we call on face processing. We propose as a hypothesis that identification of faces and categorization may follow somewhat different laws. First, though, we need to sketch in, in a mostly qualitative fashion, the type of mathematics that are applicable. Clearly, there is insufficient space for any degree of detail.
We start with the non-controversial concept of mappings (loosely, functions) between spaces to represent various levels of perceptual and cognitive processing. What these mappings can preserve, distort, etc., depends not only on the domain and range spaces, but the properties of the map, of course. Hence immediately we see that even the entire image (or more abstract representation as the case may be) is a pale simulacrum of the original object in the real world. That is, even at a relatively macroscopic level, much information must be lost in energy transduction of the narrow band and type of energy impacting the sensorium, not to mention the coarser and finer grained depictions (e.g., molecular and atomic description) that are unavailable to our perceptual mechanisms.

We then, again for most folks in an indisputable way, view a feature of an object as a map of the original object to some space or subspace. In this sense, the global perceptual object is simply the maximal feature representable by the perceiver.

Now we can get a bit more specific and take up the case of face perception. We must naturally ignore the great number of interesting theories and experiments in this burgeoning sphere (e.g., Bruce, 1988; Bruce & Humphreys, 1994; Tanaka & Farah, 1991) in order to focus on our objective. Suppose that faces, or more generally heads, as perceptual objects are manifolds (to avoid pedantry, we will drop constant reference to the original objects vs. their perceptual counterparts except where necessary). What are manifolds? Basically, they are spaces, which locally (e.g., including a neighborhood around an arbitrary point) are Euclidean, that is, can be mapped in a 1-1, bicontinuous and into a region of a finite dimensional Euclidean vector space. With a bit more structure, the differential and integral calculus can be brought into the picture and we then speak of differential manifolds and differential geometry. (For those interested, the difference between differential geometry and differential topology is that the latter requires, and utilizes geometric topics such as angle between vectors, whereas the latter depends only on more general spatial aspects; but there is much cross-talk across the areas.)

That a head, as original object and provisionally as perceptual representation can be pictured as a manifold follows immediately from approximating it by a continuous distortion of a sphere (ignoring, if their are any algebraic topologists in the audience, the “tear” distortions created by holes; not a crucial distortion of the facts for present purposes).

Now, a fair way into any decent text on differential geometry one encounters the concept of the first fundamental form (in a book on tensors, usually earlier than that). A “form” here is a kind of map that carries vectors into the real numbers, in a sense a kind of measure (e.g., feature?) on the manifold. There are several ways in which to conceive of this form. For our purposes, perhaps the most direct approach is to think of it as a means of computing lengths of paths on the manifold (which of course may be very non-flat). Recall that even in ordinary calculus, we treat velocity as a vector and calculate path length as an integral of that velocity over a path. In general, we first have to map a local region of our manifold into a Euclidean space, if we want to take an approach that includes specification of coordinates (it is possible to carry out derivations without these in most cases). However, certainly a head can be embedded (in a technical as well as colloquial sense), in a three-dimensional space, so we can just use those coordinates.

In differential manifolds, we can also view velocity as a vector, in this case as a three-dimensional vector, but how we convert this to a metric that takes into account the special properties of our particular manifold is where the first fundamental form comes in. Basically, we have this form written as

$$ f_1 = g_{ij} \frac{\partial x_i}{\partial t} \frac{\partial x_j}{\partial t} $$

where it is plain to see that the velocities play an important role.

The neat thing, though, is the $g_{ij}$, which tells how the space is changing as we move from point to point. (Actually, according to convention, the $\frac{\partial}{\partial t}$s are usually omitted, since we can integrate across the space without explicit account of them, as in line integration in advance calculus). This first fundamental form not only delivers path lengths, it virtually completely determines the so-called intrinsic characteristics of a manifold. For instance, if we lived on a planet which was forever fog-enshrouded and assuming we had no space exploration at our finger tips, all that we could ever determine about the surface aspects would be specified by this form. Nevertheless, this very powerful feature is not sufficient to explain face identification. For instance, a sheet of paper laying flat on a table and the same sheet of paper rolled into a cylinder possess the same first fundamental form; yet, we can tell the difference!

Neglecting for the time being the problem of just noticeable differences as well as obvious constancies such as translation and dilation invariances, we can ask what would be required for complete geometric congruence (in the same precise sense that we learned in high-school geometry). The answer is the second fundamental form. We haven’t much space to explicate this form, but basically, it is an extrinsic function, therefore one that is not completely determined by the first fundamental form, or map (incidentally also one that maps pairs of vectors as does the first form) that measures how any vector on the manifold is changing as we move along a path. In this sense, it is measuring something like our concept of acceleration which is itself fundamentally related to curvature of the manifold. Its expression is

$$ f_2 = f_{ij} \frac{\partial x_i}{\partial t} \frac{\partial x_j}{\partial t} $$

which looks rather like the first fundamental form. However, $f_{ij}$ is gauging how a vector that is orthogonal (i.e., normal”) to the surface of the manifold (the latter being an extrinsic, not
an intrinsic, notion) is varying as one move in different directions away from
the current position.

Anyhow, agreement of these two forms together is sufficient to ensure
congruence. Now of course, these are only two of a very large number of
structurally equivalent) encoding possibilities. Nevertheless, they do permit the
description of shapes, or much more abstract spatial objects, in a way not
imprisoned by finite dimensional vectors (although naturally the latter play a
dynamic role in the overall proceedings of differential geometry), and in a way
that completely specifies an object. We also know, of course, that many
perceptual studies have been accomplished that demonstrate deficiencies of
observers and therefore an ideal models of this sort, in their ability to
veridically perceive a number of aspects of three-dimensional space (e.g.,
Braunstein, Liter, & Tittle, 1993; Todd, 1989). Nevertheless, we believe to make
progress in object and face cognition, that we must work with the tools at hand.
As time goes by, the experimental confines can be built into emerging models,
possibly employing new or at least different geometric and topological
concepts.

We plan to begin experimentation implementing and testing models based on
differential geometry and other relatively novel mathematical objects such as
function spaces and relevant metrics on those spaces. One goal will be to start
determining just-noticeable-differences and other measures of resolution, using
such notions, in face perception.

References

shape from perspective translations and orthographic rotations. *Journal of
Bruce, V., & Humphreys, G. (1994). *Object and face recognition*. Hillsdale, NJ:
Erlbaum.
inversion effect: Testing a theory of face perception. *Perception and Psychophysics*,
20, 367-372.
Todd, J. T., & Reichel, F. D. (1989). Ordinal structure in the visual perception

Townsend, J. T., & Thomas, R. D. (1993). On the need for a general quantitative
teach of pattern similarity. In S. C. Masin (Ed.), *Foundations in perceptual theory:

Abstract

The power law and its exponent are explained by following two
alternative routes. The first model postulates a lognormal
distribution of neuronal thresholds along a dimension that data
defines stimulus intensity. Populations of sensory neurons, in the
aggregate to produce a response distribution for each intensity.
Variations in the mean and standard deviation of the lognormal
reflects the exponents of the power function for different
attributes. The second model postulates a monotonically related
stimulus intensity and response magnitude. The shape of the
psychophysical relation is determined by subjects’ preferred
number biases. Designation of the appropriate maximum response,
together with the stimulus range, determines the exponent of the
power function for different attributes.

Introduction

Physiological theory has separated into two main camps. Either
theorists treat empirical laws as due to sensory processes [8]
or they assume that results are based on conscious and/or-cognitive
varieties associated with particular scaling methods [5]. In contrast
to this duality, the Complementarity Theory of psychophysics
accepts both of these views, in an effort to explain all the relevant
data [1]. For purposes of this paper, the theory is expressed in terms
of two distinct models: the Neural Aggregate Model and the Number
Preference Model.