Available online at www.sciencedirect.com
SCIENCE DIRECT®
Consciousness and Cognition xxx (2004) xxx–xxx
Consciousness and Cognition
www.elsevier.com/locate/concog

On time, memory and dynamic form

Stephen E. Robbins

Center for Advanced Product Engineering, Metavante Corporation, 10850 West Park Place, Milwaukee, WI 53224, USA

Received 8 December 2003

Abstract

A common approach to explaining the perception of form is through the use of static features. The weakness of this approach points naturally to dynamic definitions of form. Considering dynamical form, however, leads inevitably to the need to explain how events are perceived as time-extended—a problem with primacy over that even of qualia. Optic flow models, energy models, models reliant on a rigidity constraint are examined. The reliance of these models on the instantaneous specification of form at an instant, *t*, or across a series of such instants forces the consideration of the primary memory supporting both the perception of time-extended events and the time-extension of consciousness. This cannot be reduced to an integration over space and time. The difficulty of defining the basis for this memory is highlighted in considerations of dynamic form in relation to scales of time. Ultimately, the possibility is raised that psychology must follow physics in a more profound approach to time and motion.

1. Introduction

Perhaps there is nothing more comfortably stable, more understandably a static form, than our standard "cube." The information defining this perceptual form, it seems naturally clear, must itself be a composition of static elements—"features" as we like to call them. There are problems though, as Hummel and Biederman (1992) pointed out. The cube-and-cone object of Fig. 1 has been parsed into "features" using vertices. If the same set of vertices is encountered again,

E-mail address: stephen.robbins@metavante.com.

1053-8100/$ - see front matter
doi:10.1016/j.concog.2004.07.006

YCCOG 705
5 August 2004 Disk Used
ARTICLE IN PRESS
No. of Pages 27, DTD=5.0.1
Jaya (CE) / Selvi (TE)

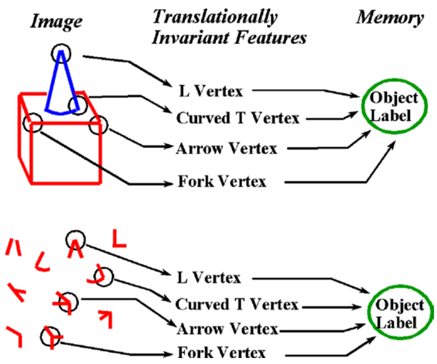

Fig. 1. Both "objects" have the same features, and therefore would be "recognized" on the basis of a feature match (after Hummel & Biederman, 1992).

presumably the object would be labeled as a "cube-and-cone" and thus recognized. Unfortunately, as the figure shows, the same set of features can be arbitrarily arranged, and thus a jumbled set would yet be "recognized" as a cube-and-cone. Groupings of more complex features yet suffer the same problem, even up to the use of 3-D features.

Still pursuing the static feature approach to form definition, Hummel and Biederman (1992) proposed that a visual object is constructed of primitive volumes termed "geons." These are elementary solids such as cones, cylinders, wedges, or bricks which are recognized by features such as straight or curved contours, straight or curved cross-sections (pyramid vs. cone), parallel or non-parallel sides (brick vs. wedge). A multi-layered (theoretical) network of feature detecting cells was described that parsed an object such as that in Fig. 1 into these elementary geons, accompanied by a description of their relationship, e.g., cone-(on-top-of)-brick. A key element in the model is the fact that a set of detecting cells is phase-locked such that their responses can be interpreted by a higher processing cell layer as coming from a single geon. The phase locking is facilitated by what is termed "fast enabling links"—high speed connections between the feature detectors.

There are multiple problems with the model, many noted by its authors, one of which is that there is no known neural mechanism for links at the velocity required. Hypotheses in which the separated brain centers processing the various features are "linked" by synchronous oscillations (Eckhorn et al., 1988; Singer, 1998), are highly tentative. How the brain uses these oscillations, or solves problems such as correlating properly the features of three cubes sitting side-by-side, is unclear. Also, similarly to the jumbled features cases already noted, arrangements of geons can be constructed that fool the system.

The geon theory is cited offhandedly in consciousness literature in the context of the processes underlying the recognition and conscious perception of forms. In general, however, the feature approach to form perception, though apparently reasonable, has been resistant to success. Similar problems have been shown to hold in machine recognition of handwriting, where feature-analytic approaches have again been tried with limited success (Freyd, 1987). Again, features have consisted of attributes like symmetry in a letter, presence of a diagonal line, open or closed curves.

Reinforcing the reasonability of the feature approach are discoveries of cells, as per Hubel and Wiesel (1959, 1978), devoted to the detection of given features—edge detectors, slanted line detectors, etc. However, additional findings demonstrated that things are more complex, for it appears rather that it is the cells' *overall* functional activity that is important. Only when all conditions are exactly the same does the detection response remain the same, otherwise cells change their response due to the external situation and due to changes in the overall pattern of the brain (Rowe & Stone, 1980). Also, the early visual receptive fields are narrowly tuned to spatial frequency. As such they are very much less selective to wide-band stimuli such as lines or bars (DeValois, Albrecht, & Thorell, 1978). The classic fields of Hubel or Wiesel, with their implication of detected features, simply cannot be regarded as the building blocks of a scene (cf. Nakayama, 1998).

The continued weakness of the feature approach leads us to the importance of the dynamic in the definition of form. The static cube-and-cone is simply a special case of the perception of time-extended events—"rotating" cubes for example, or "buzzing" flies. There are two questions here: how form might be extracted from dynamical change, and how a dynamical change is perceived as a whole event, an event with some time-extent. Neither is well understood nor seen as an underlying implication of the other. There has been a great deal of research on the former, but it is the latter, the conscious registration of a changing form over time, on which I wish to focus. The dynamic field of research on form has quite recently arrived at a crucial point, a juncture with deep implications for a theory of consciousness. It has come face to face with the problem of time. I will be abstracting the implications of our current thoughts on dynamic form insofar as these display the problem that yet faces us with respect to the perception of a whole, changing event that includes the ongoing specification of form, particularly those that relate to the concept of a "primary memory" supporting our consciousness of time-extended events.

The primacy of this problem for the question of consciousness is perhaps not appreciated. It precedes the problems of qualia. There is no dynamic form that is not simultaneously a quality. A swiftly rotating cube, almost a blur, is a certain quality; a slowly rotating cube is another. In turn, there is no such quale that has not dynamic time-extent; it does not exist in an abstract "instant" of time. Any such quality must endure or extend and change over at least two or more such "instants," else it is no quale at all. This is the most elemental condition of conscious experience. But then we quickly enter the problem of memory. If the first instant is instantly past, where, according to our standard conception, the past is equated with non-existence, we must immediately explain how the past instant is preserved and "connected" to the second, and so on, and connected such that we ultimately perceive a "rotating" cube. There will be a temptation to explain this as an "integration" over space and time, over a series of infinitesimal instants. But we may integrate over a multitude of instants, an infinity if desired, and create a 4-D structure arbitrarily large. The pure mathematical operation per se is not explanatory; it only conceals the critical problem of explaining how the first past instant endures and is prolonged into the second. We yet have the physics to explain. Immediately then, the fundamental problem of the endurance or extent of any perceived event, assuming every event has a qualitative aspect and every conscious experience has a time-extent, plunges us into the problem of the most elemental or "primary" memory of all.

I am appropriating this term, primary memory, for there is a more fundamental sense than that in which James (1890) used it, in his case, describing a memory of an event *just* past (pp. 643–645), as in the memory-image of a clock chime just finished striking. This (Jamesian sense) is something nearer our current concept of "working" memory (Baddeley, 1986).[1] Rather, primary memory underlies the very possibility of an event itself—the chime as it rings, the cube as it turns, the violin note as it sings, the fly as it "buzzes" by. It is the memory supporting the elementary perception of dynamic events or dynamic form, or, as James did in fact refer to it—our very *intuition* of time (where time is the flow of concrete events). It is within a review then of the problem of dynamic form that we approach this problem of (truly) primary memory.

2. Dynamic form

It is well known that the human visual system is capable of extracting 3-D shape from transformations of a 2-D image. This was first investigated systematically via the use of shadow projections by Wallach and O'Connell (1953), who coined the term "kinetic depth effect" to describe the perception of 3-D structure from motion information (see Braunstein, 1976; Ullman, 1979a; Todd, 1995; for reviews).

In recovering 3-D structure from transformations of a 2-D image, one is faced with an inherent ambiguity, a difficulty lying under the general heading of the inverse projection problem (see Epstein, 1995; for an overview). Unless constraints are imposed, the information is insufficient to determine 3-D structure uniquely. Chief among possible constraints is an assumed rigidity of the transforming object. Given this crucial constraint, Ullman (1979b) offered a proof that given four non-coplanar points and three successive views, a unique 3-D solution is guaranteed. If a rotating cube, for example, maintains its rigidity, Ullman's algorithm is successful.

Whether the brain lives under this rigidity constraint has been debatable. Much earlier, von Hornbostel (1922, cf. Hochberg, 1998) had provided a pertinent demonstration (Fig. 2). The demonstration involves a Necker cube slowly rotating. As is well known, when one gazes at a stationary Necker cube, the cube spontaneously reverses. In von Hornbostel's case, with the cube rotating, the reversal is accompanied by a simultaneous non-rigid deformation or distortion of the cube. Apparently, the brain's adherence to a rigidity constraint is not a constant.

Another demonstration of non-rigidity is the "rubber pencil" illusion (Pomerantz, 1983). The "rubber pencil" can be produced by hand. The pencil is held off-center, the center of gravity some distance from the point of grasp, and the pencil made to rapidly oscillate with both a translational (up and down) and a rotational (rocking) component. When executed properly, the pencil seems to become rubbery and to flex conspicuously as it moves. Pomerantz felt the effect was actually a result of persistence (or smearing) in the early stages of the visual system, and probably not in the later stages where Ullman's rigidity constraint is presumably used in computations. However, he

[1] Neither is this "sensory" memory, "iconic" or "echoic" memory. The interest here is not in a memory that preserves several "items" of "sensory information" for less than a second, during which decisions occur on what to send on to short term memory, or on which "features" to transfer to STM. This is an utterly static conception of memory. Has perception occurred, or is it occurring, while such a memory is operative? Is this memory supporting perception? Or is this truly just "sensory" information? What is the role of such a memory in the perception of a dynamic event such as a rotating cube? What if, as we shall see, static features are a fiction? These questions are unasked and unanswered.

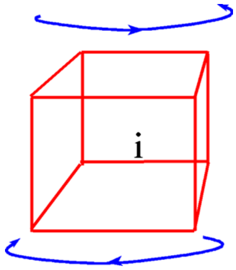

Fig. 2. Von Hornbostel's (1922) wire cube, rotating in the direction shown at the top of the cube, with (i) being the rear edge. The cube reverses spontaneously, (i) becoming the near edge, rotating as below, distorting as it does so.

felt even this called into question the generality of the constraint, for if the constraint is so powerful, why has the vision system not learned to account for smearing in its early stages?

Ullman (1984, 1986) also presented an "incremental rigidity" scheme or computational approach for extracting structure from moving objects, both rigid and non-rigid. Its key was maintaining an internal model of the object and modifying it at each instant by the minimal non-rigid change sufficient to account for the observed transformation. Thus, given an internal model, $M(t)$, and a new frame of (x, y) values, the problem is to find a vector of depth values such that the overall deviation from rigidity is minimized. Again supposing our cube as the initial object in rotation, the cubical structure will be specified gradually as the object moves by Ullman's algorithm. Now suppose the cube, while rotating, begins to transform gradually to a trapezoidal solid (as seen from the top), completing this change over the course of two full (360°) rotations. Ullman's model nevertheless "keeps up," computing in fairly close order a structure close to the actual developing form, i.e., the forming trapezoid. If the amount of the change is too great over a time, e.g., a completed transformation to a trapezoidal solid within 180°, the model is overwhelmed, its computation becoming less and less close to the actual form at any instant.[2]

Despite Ullman's argument, the status of the rigidity constraint, and therefore of these algorithms for extracting form, is less than secure. Hochberg (1998) dismisses the constraint's general validity. Many non-rigidity effects, besides that of von Hornbostel, appear not to fit.[3] And not least is the simple fact that we perceive non-rigid forms. Thus several experiments (cf. Todd, 1995) show compelling kinetic depth effects of objects moving in 3-D space, though non-rigid. The computational models utilizing variants of the rigidity hypothesis, such as piecewise rigid motions composed of locally rigid parts whose relative spatial arrangements can arbitrarily deform

[2] Another class of models (Todd, 1982; Webb & Aggarwal, 1981), applicable to rigid and non-rigid 3-D structure, uses the relative trajectories of each element's projected motion rather than their positions at discrete moments in time. These still require samples over a sufficient period of time to register the individual trajectories.

[3] Several non-rigidity cases are mentioned in Todd (1995). Another illusion first reported by Ames (1951) will be noted in a later section. Beghi, Xausa, and Zanforlin (1991) and Beghi, Xausa, De Biasio, and Zanforlin (1991) attempted to account for non-rigidity in an ellipse rotating in the image plane (and other effects) using a minimal relative motion principle, suggesting the rigidity constraint is unnecessary, rigidity being a by-product of their principle. Liu (2003) has given a definitive critique of their approach, but noted that a different minimization principle which can be achieved through local computation, e.g., the "slow and smooth" hypothesis of Weiss, Simoncelli, and Adelson (2002) (discussed below), may make headway toward explaining rigidity in percepts.

(e.g., Hoffman & Flinchbaugh, 1982; Todd, 1982), generally do not degrade gracefully. If there is no rigid interpretation, they cannot say anything at all about the object's 3-D structure (cf. Todd, 1995).

The tale of the rigidity constraint is simply an illustrative one in the problem of the relation of time to form. Ullman (1984) clearly saw the critical role of time in deriving 3-D structure. He noted that Wallace and O'Connell included 3-D objects constructed of wire whose static projection produced no 3-D perception. These objects appear initially flat when viewed in the frontal image plane, but acquire their correct 3-D structure when put into motion. He rejected models that require the storage of long sequences of different views or samples, yet, he argued, clearly the system has to be able to integrate information over an extended viewing period. But while incorporating this requirement, his internal model, $M(t)$, as earlier noted, rests upon the specification of structure found within it at any "instant." It is the extent of this "instant" that is a crucial, yet unexamined component of any approach, whether or not storing samples of an event.

2.1. The Shaw–McIntyre demonstration

Let us consider now a demonstration discussed by Shaw and McIntyre (1974). In the demonstration, we have a cube constructed of wire edges and rotating at a constant speed. Every such object has a symmetry period. If we consider rotational symmetry, the period is given by the number of times the object is mapped onto itself or carried into itself in a complete rotation of 360°. Thus a square, if rotated about its center, is completely carried into itself every 90° and has therefore a symmetry period of four ($4 \times 90 = 360$). An equilateral triangle has a symmetry period of three, being carried into itself every 120° turn ($3 \times 120 = 360$). A circle is considered to have a period of infinity, being completely carried into itself with even the smallest rotation.

If the room is dark and we strobe the cube periodically, the form that is actually perceived is totally dependent on whether or not the periodic strobes preserve this symmetry (invariance) information! If we strobe in phase with or at an integral multiple of this period, an observer would see, as we might expect, a cube in rotation (Fig. 3). But if the strobe is out of phase, what is perceived is not a rigid cube in rotation, but *a distorted*, *wobbly object*.

Let us explore the significance of this demonstration. Consider firstly the implications for a model like Ullman's. In this model, the form of the cube can be found at any time, t, via the model

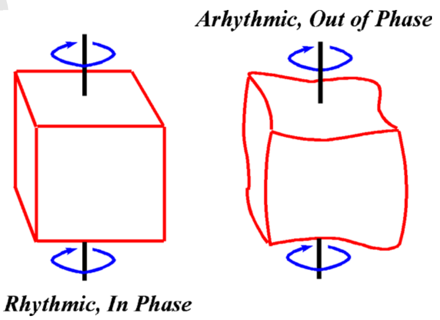

Fig. 3. Rotating cubes, strobed in phase with, or out of phase with, the symmetry period.

$M(t)$. Though Ullman's (1984) algorithm can be implemented continuously, it is perfectly happy with a succession of samples, and works with the samples independently of any phase relationship with the object's symmetry period, i.e., the arrhythmic sampling would yet provide samples sufficient for the algorithm to compute both a cubical and a rigid form. Given Ullman's principle of three successive views, then at $t_1 + 2$ the form of the cube—as a rigid cube—should be specified, should subsequently stay specified, and be found at any instant, t ($t > 2$), in $M(t)$.

What is important for Ullman's algorithm, what is "specific to" form, is the information computed in $M(t)$ at the instant, t. However, Shaw and McIntrye's demonstration tells another story. The form of the cube, even the rigidity of the cube, is actually an invariance defined over time. The successive self-mappings define a periodicity, something we could view as a sinusoidal wave. But this is something an arrhythmic strobe does not specify. It is arguably this implicit periodicity that specifies the form. It would be this pattern, over time, that is critical. Such an invariant cannot exist in the instant, t.

2.2. Structure from optic flow

Gibson (1950, 1966) insisted on the importance of optic flow fields in specifying depth (Fig. 4). Models (e.g., Domini, Vuong, & Caudek, 2002; Hildreth, Ando, Andersen, & Treue, 1995) that extract 3-D structure from optic flow are yet aligned with Ullman, both with respect to the rigidity constraint and to a 3-D representation whose accuracy builds up over time, but the rigidity constraint has recently been completely abandoned (Domini & Caudek, 2003).

Consider a rigid flag, rotating around a pole. As the flag rotates away from the observer (or away from the frontal plane), the resulting optic flow is characterized by a pure contraction. At the flag rotates towards the observer, the optic flow is a pure expansion. Given that this flow is constant, the magnitude of the perceived slant, according to current models (see Domini et al., 2002) depends solely on the intensity of the flow gradient at t_0. The perceived slant at any time should then be constant, but in reality, to the human perceiver the surface slant appears to be continuously increasing in the course of time as a consequence of the perceived rotation.

To account for this, Domini et al. (2002) offer a model requiring temporal integration on a longer scale. The (short-term) time required, Δt, for the visual system to measure the intensity of the gradient, is by consensus approximately 150–200 ms. The Domini et al. model simply

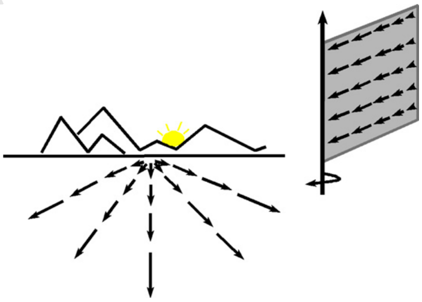

Fig. 4. Optical flow field. A gradient of velocity vectors is created as an observer moves towards the mountains. The flow field "expands" as the observer moves. At right, the flow as a flag rotates towards the observer.

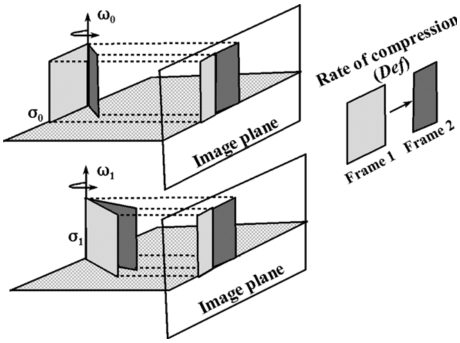

Fig. 5. Image transformations via rotation. Initially (top, and also Frame 1) the surface (e.g., a stiff flag) projects a square (light gray) on the image plane. This is compressed after rotation (dark, Frame 2). The rate of compression is the amount of def. It is measured by considering the value of the velocity vectors in two successive instants of time. The same compression (bottom) can be created by an initially larger slant and a small rotation (adapted from Domini & Caudek, 2003).

hypothesizes that the perceived slant is a function of the weighted average of the slant magnitudes derived at moments t_i and t_{i-1}. The model (which can be stated in continuous form) successfully predicts perceived slant. Still, though it requires time to get there, the specification of the form or 3-D representation (slant) in this model ultimately lies in the instant, t.

Local optic flow (as in a patch of a surface) is inherently ambiguous. Consider the optic flow created by the rotation about the axis (ω) of a planar surface, again, a stiff flag (Fig. 5). After the rotation, the initial projection on the image plane will be compressed. The compression is expressed by a quantity termed def, where def expresses the result of either of two components of shearing that change the shape of the projected surface (def = $\sigma\omega$). The same compression produced by a small slant (σ) and a large rotation can be created by the small rotation of a more slanted surface. There are an infinite number of such (ω, σ) pairs. If the amount of perceived rotation for each surface of a form, e.g., a cube or a wedge, is derived from def, then, in general, due to this ambiguity, reliable discrimination between rigid and non-rigid motion is not possible, and this has been demonstrated (cf. Domini & Caudek, 2003). This is an accord with the conclusion that neither Euclidean nor affine properties are recovered veridically from optic flow, therefore algorithms based on these properties are insufficient to the task.[4]

Domini et al. (2003) have thus been led to propose a probabilistic (Bayesian) approach to determining form, wherein observers interpret local optic flow by choosing the most likely solution maximizing the posterior probability of a σ, ω pair given def. Rather than solving the inverse projection problem, the system proceeds in a probabilistic patch-way fashion via optic flow. The same

[4] Affine geometry is a "rubbersheet" geometry, concerned with preserving relationships under transformations, but not Euclidean distance. The three-frame criterion (e.g., Ullman, 1979b) is applicable only to Euclidean distance relations between pairs of points. Two-frame motion sequences yet serve to provide information for any object property invariant under affine (stretching) transformations, and numerous demonstrations show the kinetic depth effect can be created with two-frame motion sequences. Both rely on the rigidity constraint. Yet Domini and Braunstein (1998) have shown that neither Euclidian or affine algorithms suffice.

Bayesian/optic flow approach has been proposed in the context of the even more basic "energy" models which we will discuss shortly. Thus Domini and Caudek state, the fundamental problem now is understanding "how local information is integrated through space and time to achieve a global and coherent 3-D shape." As initially noted, this is simultaneously to say we must enter the problem of time and memory.

Now it will be relevant to imagine for a moment, in this context of optic flows, how one might describe a "Gibsonian" cube. If a rotating, rigid flag is described by expanding and contracting flow fields, a rotating cubical solid is a partitioned set of these flow fields, expanding and contracting as the cube's faces come into view and leave. The cube's edges (straight lines) and vertices will be sharp discontinuities at the junctures of these gradient flows. These "features"—straight lines, vertices, even "rigidity"—would be very ephemeral, dynamic things, pure creatures of time, in a Gibsonian cube.

2.3. Form and internalized laws

The probabilistic approach towards which the optic flow theorists have been led naturally induces the question of the origin of the probability estimates the system must apply to its inherently ambiguous optic flow information. The rigidity which the "Gibsonian cube" does achieve is perhaps a function of higher order constraints or laws the system employs for its probability computations. One of the more famous phenomena in the realm of higher order laws or constraints determining form is an illusion first reported by Ames (1951). When a static trapezoid is observed monocularly in the frontoparallel plane, it is typically perceived as a rectangle slanted in depth. As it begins to rotate, under appropriate conditions, the trapezoid will appear rather to oscillate, and even undergo severe non-rigid deformations.

Ames and other transactional functionalists of the period emphasized the role of experience (hence "transactions") in determining perceptions. Ittleson (1962) would speak of "assumptions" or weighted averages of experiences that are brought to bear in any given opportunity for perception. Thus any concrete experience involves the resolution of a host of possibly incompatible assumptions via some unconscious weighting process based in probabilities derived from experience. Beyond this, the transactional functionalists offered only little in terms of theory, but they were precursors of theories of higher order processes determining form.

Carlton and Shepard (1990a) echo Ittleson to an extent, arguing that the representational process arises as a kind of "resonance" to external stimulation within a system that—through natural selection and experiential fine tuning—has already internalized the principal regularities of the world. They go on to argue:

> "An essentially full representational response may be excited within the resonant system by a fractional part of the external information. That is, the missing information is automatically interpolated, extrapolated or otherwise completed by a kind of unconscious simulation guided by internalized principles about how the world works" (pp. 133–134).

We will consider now certain higher order principles Carlton and Shepard proposed to guide this extrapolation or completion, and ultimate "resonance." We will see that these principles may embrace the non-rigidity of Shaw and McIntyre's cube, but still leave the question of time-extended form.

YCCOG 705
5 August 2004 Disk Used
ARTICLE IN PRESS
No. of Pages 27, DTD = 5.0.1
Jaya (CE) / Selvi (TE)

2.4. Form and the geodesic path

Consider a cube in two successive positions/orientations (Fig. 6). Via Chasles' theorem, the motion from one position to the other is described by a screw displacement. This screw-motion consists of a rotational and a translational component. The center of the object then traverses a helical path. Carlton and Shepard note that in experiments in which objects are presented (at a brief interval) and seen in two successive positions, observers indeed report a screw-like or helical apparent motion, evidence, apparently, of an internalized rule.

The successive positions of a cube create a manifold of possible positions—a 6-D space. This 6-D space is a product of two 3-D spaces—the 3-D space of possible locations (denoted R^3) and the 3-D space of orientations of the cube (denoted SO(3)). Carlton and Shepard argue that the psychologically simplest motions from any point to any other point in this 6-D manifold are characterized by the straightest possible paths, termed *geodesics*. The geodesics are characterized by Chasles' helical or screw-like path.

If we focus on the symmetry group of the cube, it is a subgroup of the space of possible orientations, SO(3), consisting of 24 elements (rotations about axes through the faces, rotations about axes through the diagonals, and rotations about axis though opposing edges). Thus there are 24 distinct screw displacements. For any possible motion in this subspace, the motion preserving the greater symmetry is preferred. Thus, for a given motion from one orientation to another, a path corresponding to a motion about the axis of 4-fold symmetry (e.g., an axis through two opposing faces) is preferred over a motion about an axis of 3-fold symmetry, etc.

Within this manifold, the distances along geodesic paths between points A and B correspond to rigid transformations, whereas the *direct* distance between A and B corresponds to a non-rigid transformation. Non-rigidity corresponds to a shortcut through the embedding space (Carlton & Shepard, 1990b). As the geodesic paths become longer, and the shortcuts relatively shorter, the probability of taking the shortcut increases.

This leads to the essential form of explanation Carlton and Shepard (1990b) would apply to Shaw and McIntyre's non-rigid cube. Given the symmetry of a normal cube (with unlabeled faces), there are additional connecting screw motions between any two orientations. If the symmetry is only approximate, as for a nearly cubical rectangular solid, then some of these paths are excluded as rigid transformations. The alternative short-circuit paths corresponding to small

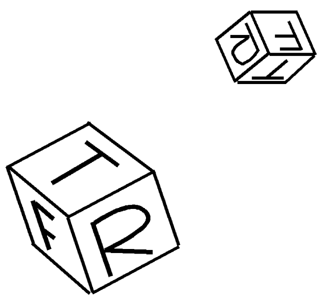

Fig. 6. A (labeled) cube in two successive positions and orientations. There are multiple paths by which the cube might rotate and/or translate from its original position and orientation to the other (adapted from Carlton & Shepard, 1990a,1990b).

non-rigid transformations are expected to increase, and decrease the probabilities of traversing the longer geodesics corresponding to rigid transformations. Thus, if the arrhythmic strobe is reducing the cube's symmetry, making it only approximate, the probability of taking these alternative, short-circuit paths corresponding to non-rigidity is increasing.

The wobble of the cube may also follow. The symmetry axes of an object are the favored axes of rotation, and for objects with a salient axis, the rotational component is the favored motion. Rather than being about an axis that is fixed in the environmental frame, Carlton and Shepard argue, this may be a rotation about an axis that is inherent in the object itself and that therefore must be represented as moving in the environmental frame. With an axis moving relative to the environmental frame, we have a wobbling object.

Assuming that the arrhythmic strobe specifies or induces an irregularity (non-symmetry), the general non-rigidity and wobbly motion of the cube is at least in conformity with Carlton and Shepard's laws. However, this begs the question of why the arrhythmic strobe induces the lack of symmetry. There is the possibility that a higher, temporal constraint exists, expressed simply: temporal regularity implies spatial symmetry. The two are reciprocal. A rotating symmetric cube "pulses" a temporal regularity via its symmetry. The wide picture window with its regular panes pulses a regularity as my glance passes across it. The arrhythmic strobe may be disrupting this fundamental law of experience or design. The discussion of the ambiguity of optic flow has already given a glimpse that all form is a function of probabilities. The probability estimates must come from higher constraints. The very rigidity and structure of the rotating Gibsonian cube would then be literally held together via this yet higher, but disruptable, temporal symmetry constraint.[5]

This leads us to the most detailed approach to these probabilistic constraints.

2.5. *Form—the approach from below*

Adelson and Bergen (1985) described a general class of low-level models based on linear filters known as "energy models," initially developed by Watson and Ahumada (1983), for detecting the elements of dynamic form. These are addressed specifically to the detection of the direction and velocity of motion, for example, as an edge of our cube transits the visual field. They are an evolution from the correlation filter (Fig. 7) of Reichardt (1959) for motion and speed detection, and there are significant formal connections.

Adelson and Bergen discuss the energy model in the context of the apparent motion specified in perception when presented successive frames at brief intervals, as in a movie. It specifically bypasses what is termed the "correspondence" problem, and simultaneously, a need to track changing "features." The correspondence problem is generated by considering the possibility that the visual system matches corresponding points or features in successive frames, determines Δx, the

[5] Von Hornbostel's distorting, non-rigid, rotating Necker cube itself demands explanation. A static Necker cube can be projected onto two different 2-D topological mappings (Shaw & Mace, in press), between which the brain apparently oscillates. But why such oscillation (in the midst of the object's rotation) might bring the cube under Carlton and Shepard's geodesic law is not immediately clear. It could be noted that the Shaw and McIntyre cube is a rotating Necker cube as well, but Shaw and McIntryre do not report it as distorting under the rhythmic strobe as it rotates—an interesting difference.

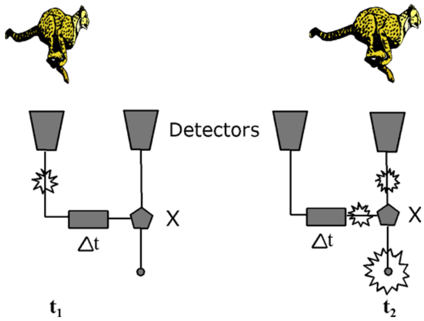

Fig. 7. Reichardt filter or correlation model (Reichardt, 1959). It has two spatially separate detectors. The output of one of the detectors is delayed and then the two signals are multiplied. The output is tuned to speed. Many detectors tuned to different speeds are required for the true speed of a pattern, and the difference of pairs of detectors tuned to different directions is taken.

distance traveled, and Δt, the time between frames, and computes, $v = \Delta x/\Delta t$. These matching models, particularly given the question of what constitutes a "feature," prove to be intractable.

The energy model does not extract position to compute motion. Motion is treated as spatiotemporal orientation (Fig. 8), and the model consists of a network of spatiotemporal filters. The response of the spatial component of the filter is the sum of its responses to varying local intensities of light falling in its receptive area, point by point. The key here is to think of the temporal response of these filters as a temporal weighting function which describes (as above in Domini et al.) how inputs in the past are summed to produce the response at the present moment. The filters thus respond to motion energy within particular spatiotemporal frequency bands (hence these are also termed Fourier models, since the space or coordinate structure is spatial and temporal frequency). A network of these filters distributed across the visual field produces a net form of continuous output specifying the direction and velocity of motion of the edge.

Adelson and Bergen's filters will account for a phenomenon such as reverse phi motion, where a grating of bars (alternating black and white) in motion in successive steps to the right (in actuality) will be perceived as moving to the left if the black and white bars are interchanged at each step. The model shows, in fact, a predominance of leftward-moving energy in this case. There are, of course, cases of ϕ motion that have typically been taken to indicate higher order laws at work. For example, a dot, flashed in two successive positions, first on one side of a square, then the other

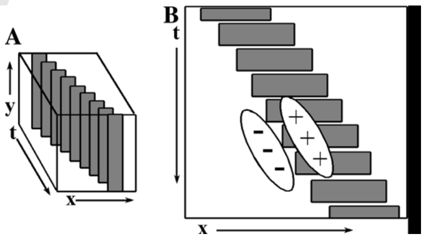

Fig. 8. Motion as orientation in (x, t). (A) Spatiotemporal picture of a moving bar sampled in time. Velocity is proportional to the slant. (B) Spatiotemporally oriented receptive field that could detect the bar's motion (adapted from Adelson & Bergen, 1985).

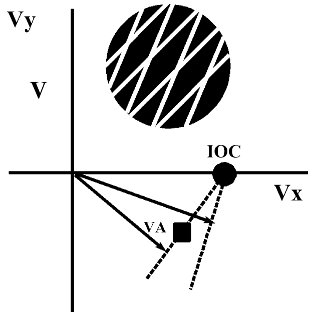

Fig. 9. The plaid pattern is the result of the motion of two gratings. Each grating has a velocity vector normal to the grating lines, lying on a constraint line in velocity space. Vector averaging (VA) takes the average of the two normal vectors. Intersection of constraints (IOC) finds the single velocity consistent with both sources of information (after Weiss & Adelson, 1998).

side, appears to leap or go behind the intervening square (cf. Hoffman, 1998; for examples). However, the installation of simple constraints in these energy models is already moving, as we shall now see, towards the explanation of a large set of motion phenomena.

Perhaps at the current apex of the movement from below towards the achievement of the perception of global form is the model of Weiss et al. (2002). A piece of background is yet in order. Consider the "plaid" grating of Fig. 9. This is composed of two oriented gratings crossing each other in the image plane, and indeed, when viewed separately, each grating is seen traveling in its oriented direction. Yet when the gratings are presented simultaneously, we see them moving coherently, assigning a single motion to the pattern. This phenomenon is a function of the "aperture" problem, a problem created when the ends of the lines are not visible.[6] The individual velocity measurements provide only a partial constraint. Considering each grating, only the component of velocity normal to the orientation of the grating can be estimated, and hence the grating motion is consistent with an infinite number of possible velocities—a constraint line in velocity space. The visual system, it was discovered, depending on various conditions, appears to resolve this by either of two possible rules (Intersection of Constraints or Vector Averaging).

It should be remembered that the receptive fields of the energy model filters discussed earlier are inherently "apertures." The aperture problem indicates that the visual system's measures of velocity are intrinsically *uncertain*. Therefore the integration of a multitude of uncertain individual velocities must be inherently probabilistic. It is at this point of integration that Weiss et al. insert their fundamental, Bayesian constraint.

Bayesian inference contains these basic elements—the combination of estimates while factoring in their uncertainty, and the integration of prior knowledge during the combination. If I am trying to determine if my favorite football team is presently posting a certain score in the third quarter, I can ask two friends (A and B) for their estimate of the score, asking each to provide an estimate of

[6] Imagine the slanted lines of one of the gratings composing the plaid pattern of Fig. 8 as very long, in fact on a large card. The card is moving directly to the right and the passing lines are seen in the window of the circular aperture. The component of velocity moving to the right is not seen. Only the component of the velocity orthogonal to or normal to the lines is seen. This component is moving downwards.

YCCOG 705
5 August 2004 Disk Used
ARTICLE IN PRESS
No. of Pages 27, DTD = 5.0.1
Jaya (CE) / Selvi (TE)

the certainty or likelihood of his opinion (E_a and E_b, respectively). I myself have a prior knowledge of the possible score (S). In Bayes formula this becomes, assuming each friend's knowledge is independent of the other:

$$P(S|E_a, E_b) = kP(S)P(S|E_a)P(S|E_b).$$

This expresses the posterior probability of the current score as a function of the likelihoods and prior knowledge. (Here k is a normalizing constant independent of S.) In analogous fashion the first stage of the model uses the output of spatiotemporal filters to gather motion information from each small image patch. The measurements are used to obtain a local likelihood map, where for any particular candidate velocity the likelihood of the spatiotemporal data generating that velocity is estimated. The prior knowledge or assumption that Weiss et al. have employed is this: motion is slow and smooth. The observation that humans tend to choose the "shortest path" has been supported since the turn of the century. (We have already seen a form of it in Carlton and Shepard's geodesics). As the bias for slowness would lead to highly non-rigid motions in curved figures (Fig. 10), the smoothness constraint (e.g., Hildreth, 1983) has been suggested, where adja-

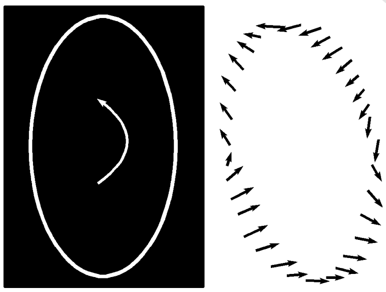

Fig. 10. The normal velocity vector components (right) of the edge of the rotating ellipse (left). These tend to induce non-rigid motion (after Weiss & Adelson, 1998).

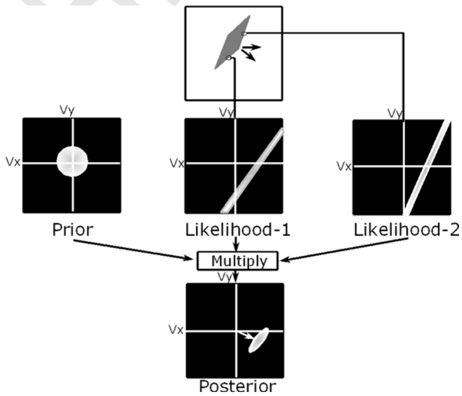

Fig. 11. The low-contrast rhombus (top), when moved to the right, is actually seen to move downward (lower arrow). The probability distributions of velocities, falling along constraint lines in velocity space, are shown in the middle row. The *prior* probability of velocities in (x, y) is multiplied together with the *likelihoods* of an edge moving at each velocity, obtaining the *posterior* (after Weiss & Adelson, 1998).

cent locations in the image have similar velocities. This joint constraint is expressed as a prior probability on velocity fields that penalizes or defines a ‘‘cost’’ for the speed of the velocities and the magnitude of their derivatives. Velocity fields corresponding to rigid translation in the image plane, for example, will have a high probability—since velocity is constant everywhere as a function of space, the derivatives will be zero.

The model demonstrates that a large number of illusions of motion can be explained with this Bayesian approach (cf. Weiss & Adelson, 1998; for a more extensive list). One example is shown in Fig. 11. To note another, there is the illusion observed by Mussati (1924) and Wallach, Weiss, and Adams (1956). A narrow ellipse rotating in the image plane deforms non-rigidly, but if the ellipse is ‘‘fat,’’ rigidity in the percept becomes prominent. For the narrow ellipse, due to the prior which favors smooth and slow velocities, the estimate is biased away from veridical velocity, and toward the normal velocity components (Fig. 10), moving then to non-rigidity in the percept.

This extremely powerful and dynamic body of thought is apparently moving towards incorporating the constraints the transactionalists envisioned. How complex the Bayesian priors become remain to be seen. It is possible that the geodesic constraint discussed above can be subsumed when the model is taken to the 3-D case.[7] The framework contains certain deep lessons. One is the intrinsic embedding of uncertainty within the system. This is framed in the context of obtaining velocities given the aperture problem. But, we shall see, there must equally be uncertainty in obtaining velocities due to the dynamic motion of time itself, for there is no such thing as a static instant such that a velocity could ever be completely determined. The entire system therefore is embedded in an uncertainty that is a function of the matter-field in which it is embedded. As Weiss et al. insist, these ‘‘illusions’’ are percepts precisely as optimal as they can be given the information available. Second, form itself has clearly become a creature of time, a function of velocity fields. Even the specification of rigidity is a consequence of constraints met, or not met, by these moving fields. This is the death of the notion of static features. Finally, for all the power of this model, there is yet no perception; there is no possibility of the *experience* of dynamic form. A reflection on its origins in the Reichardt filter (Fig. 7) initiates the point. The past instant on the trajectory of a moving object is registered, a delayed signal generated, and then multiplied by a signal from a following instant on this trajectory. The multiplied signal value, at this new instant, is taken (assuming a network effecting some disambiguation) as indicative of direction and velocity. It reflects an influence of the past. But it is simply another instant. The network vectors towards this output, instant after instant. The filter (or an entire network of spatiotemporal filters) will *register* direction and velocity, i.e., produce a value, or successive values, but though the successive signals are correlated, this is not the *perception* of a motion. It does not fulfill the elementary criterion for consciousness asked for in the introduction—a continuity over at least two such instants. There is no experience. As external observers, we see the energy model, and the powerful computing network it embodies, computing values indicative of, and a basis for,

[7] We can ask where the geodesic constraint will fit as the Weiss et al. model is generalized beyond 2-D. A geodesic path that is long relative to the time between the successive positions of a cube would require a very fast motion, tending to violate ‘‘slow and smooth.’’ A slower motion would involve a deformation but is acceptable under this constraint. It would seem, however, that the geodesic constraint has a hierarchical priority. Shepard implies that the deformation (under slow and smooth) would actually be preferred, yet is resisted, the geodesic predominating—until the discrepancy in path length vs. shortcut through the embedding space is too large.

YCCOG 705
5 August 2004 Disk Used
ARTICLE IN PRESS
No. of Pages 27, DTD = 5.0.1
Jaya (CE) / Selvi (TE)

a changing form over time. But it is we, as external viewers, that are assigning this continuity. From an internal view, there is no such continuity; the fundamental cohesion of time is missing. Ultimately, as noted in the introduction, this places us in the heart of the hard problem, and equally in the heart of the question of the additional ingredient required for a consciously perceiving device. Let us consider time, memory and form then more deeply.

2.6. Form and the scale of time

The cube, rotating at a certain rate, and perceived as a cube in rotation, is a function of a scale of time imposed by the dynamics of the brain. We could increase the velocity of the cube's rotation. With sufficient increase, it will become a serrated-edged figure, and at a higher rate, a figure with even more serrations. Finally, it becomes a cylinder surrounded by a fuzzy haze. Each of these figures is a figure of $4n$-fold symmetry—8-edged, 12-edged, 16-edged,. . ., with the cylinder a figure of infinite symmetry (Fig. 12). In total, this transitional series of forms reflects the scale of time in which we normally dwell.[8] Beneath the specification of these forms, we can posit, lies an attractor, supported by the brain's dynamics. Let us perform a gedanken experiment. Underlying this dynamics are chemical velocities supporting the brain's computations. The range and complexity of these, considering the various local velocities, is vast, but at least in principle, it can be argued (cf. Fischer, 1966; Hoaglund, 1966), the global process velocity could be changed; we could introduce some catalyst or set of catalysts to effect this.

Suppose then two observers. Observer A, dwelling in our normal scale, is gazing upon a cube rotating rapidly enough to be perceived as a 16-edged serrated figure. Observer B has had his global process velocity raised. His scale has been shifted. He perceives the same cube as a cube of normal 4-sided construction slowly rotating. Both perceive by the same law of invariance—a figure of $4n$-fold symmetry. Suppose A and B are watching a time-lapse film of the growth of a human head in profile (Fig. 13). At a given film rate, the head is transforming very rapidly for A; for B, it is a much slower event. Both perceive the transforming head by the same law—a strain transformation applied to a cardioid. Were we to borrow from physics, we might say that we have performed an operation analogous to changing the "space–time partition." From this perspective, the significance of Gibson's (1966) insistence on invariance laws defining events could be understood. In such transforming partitions, it is only invariance laws (e.g., $d = vt$, $d' = vt'$) that hold.

Extend this scale transformation. Raising the process velocity carries a corresponding decrease in the scale of time. If a fly is passing before both observers, for observer A it is a "buzzing" fly of our normal scale; for observer B, the fly is now flapping its wings slowly, like a heron. As the process velocity of B is raised further, the fly transforms, transitioning to a near-motionless fly with wings barely moving, then to a motionless fly, then to a collection of waves. . .. Each form represents an ever smaller scale, but this is equally to say that each successive scale requires a lesser degree or power of "primary memory."

[8] A structure-from-motion algorithm such as Ullman's (1979b) would fail to represent the transitional serrated-edged forms arising as the cube is rotated at increasing velocities. The constraint required for the energy model to reflect this is a question, but it should be better capable of reflecting the increasing velocity as change in form.

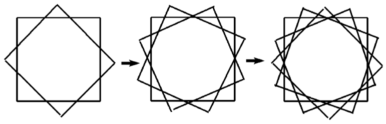

Fig. 12. Successive transformations of the rotating cube (2-D view) through figures of 4*n*-fold symmetry as angular velocity increases.

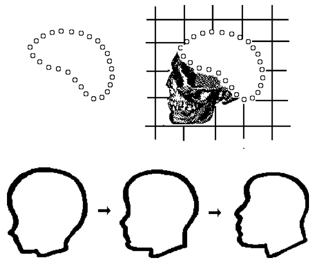

Fig. 13. Aging of the facial profile. A cardioid is fitted to the skull and a strain transformation is applied. (Strain is equivalent to the stretching of the meshes of a coordinate system in all directions.) Shown are a few of the possible profiles generated (adapted from Pittenger & Shaw, 1975).

2.7. Primary memory and dynamic form

What is primary memory? First, it is a heavily unanalyzed problem, and second, a concept intrinsically related to the "instant" of perceptual theory. Gibson (1975) stated the issue succinctly:

> The seemingly innocent hypothesis that events are perceived has radical implications that are upsetting to orthodox psychology. Assuming that shorter events are nested within longer events, that nothing is instantaneous, and that sequences are apprehended, the usual distinction between perception and memory comes into question. For where is the borderline between perceiving and remembering? Does perceiving go backwards in time? For seconds? For minutes? For hours? Where do percepts stop and begin to be memories, or, in another way of putting it, go into storage? The facts of memory are supposed to be well understood, but these questions cannot be answered. (1975, p. 299)

The "buzzing" fly, a creature of our normal scale, is a perception, perhaps better, a single visual intuition, spanning hundreds of wing oscillations per second, all summed up as a blur. The rotating cylinder with surrounding fuzzy haze, in actuality a rapidly spinning cube, is again a "blur" taken over a good number of revolutions.[9] A hand waving "good-bye," as an event, contains multiple oscillations, all perceived as a "wave." Consider the fly. Assign (arbitrarily) to an "instant," a wing-beat cycle. We could obviously be more fine-grained, assigning a fraction of an arc of wing movement to an instant, or far less. Each instant/cycle, as it occurs, is by definition "present." As

[9] It should be apparent here that *scale* implies *quality*. Simultaneously, scale implies *time-extent*. This must at least be considered as the origin of, the base support for, qualia.

the next cycle or wing-beat comes, the previous is now "past." We must then assign it to memory. To account for the perception of the "buzzing" fly, we now have a stack of stored past instants, all of which are retained and integrated or related somehow to form the whole event. A portion of this stack somehow "slides along" in time as the "present" or ongoing perception, some of the earlier instants falling off, others being added on as they arrive. This is a strange process; the logical problems, we shall see, are enormous. It relies on a memory that stores "instants" and supports the perception of events consisting of sets of instants, events clearly delimited in time, yet large enough in extent to support the perception of a whole event such as "bye-bye" wave, a "buzzing" moving fly, a "rotating" cylinder or cube. This is primary memory.

Yet, earlier, we have seen that we can treat the brain as a dynamical system that imposes a scale upon time, i.e., that imposes a scale upon the matter-field in which the brain is embedded. Further, it may well be possible, at least in principle, to change the process velocity or chemical velocities underlying the dynamics of this system, and in so doing, alter this imposed scale. With the proper catalyst(s) we could change the scale such that the buzzing fly is now nearly motionless, its wings barely "inching" downwards at increments of fractions of an arc. In other terms, the space–time partition is changed. Consider again Messrs. A and B. The razor-thin instant or "present" of the transforming universal field is the same for both. A sees the "buzzing" fly—hundreds of wing beat cycles forming A's "present." B sees the nearly immobile fly. For B, a wing-position "a few seconds of an arc ago" is in the vastly far past. The hundreds of cycles comprising the "buzzing" perceived by A as his "present" are vastly in B's past. Does B have the right to say all these are *non-existent*, being preserved only by A's primary memory? Yet we could imagine a being C, at even higher process velocity and yet smaller scale of time, arguing the same of B's minute changes of wing position.

Further, let us note that the catalyst(s) could be increased gradually in B, spreading out the "buzzing" event all along its time-spectrum, inducing a gradual transformation of the fly, from buzzing, to heron-like, to immobile, etc. Does this imply that B's primary memory (or his need for this memory) is decreasing, capable of storage of fewer and fewer "past" events? We could make the same example of the cube, for A, a rotating, fuzzy-bordered cylinder, for B a stable cube. If, for A, via our process velocity increases, the fuzzy cylinder event is gradually spread out, the event will gradually transition, punctuated by "quantum transitions" as it moves through forms of 4*n*-fold symmetry, until it reaches motionlessness, and then moves on to even smaller scale events comprising the matter-field of the cube.

This introduction of what we could term a "relativity principle" makes extremely problematic the assigning of a function to "primary memory" by which the "past" is preserved from non-existence and the perception of time-extended form is enabled. It what sense is this a "memory?" In what sense, if we can simply attribute a change of the extent of its retention of the past (the number of events stored) to a change in the underlying dynamics? In essence also, this is the introduction of the "relativity of the instant." This instant always has some thickness, some temporal extent. But the basis for the primary memory which supports these extents in perceptions and the forms of these perceptions is a far from simple concept of a "memory."

2.8. Abstract space and time

When we consider, then, our simple rotating cube and its significance, it appears that it is a dynamic pattern, defined over time, specifying the form of the cube. The "features" that define form are not the static entities, existing in some instantaneous slice of time, we have believed. As Gibson (1966, 1979) long argued, the concepts of our Euclidean geometry—straight lines, curves, vertices, sets or families of forms related by geometrical transformations (e.g., Ullman), even geons—while elegant, may have little meaning to the brain, i.e., they are not the elements by which the brain constructs a world. They do not lie in the "instant." These very concepts are constructions, projections of the brain—part of a projective framework of thought. They are an integral part of the projective thought-framework of *abstract* space and *abstract* time—a space and time in which the brain does not actually dwell.

Bergson (1896/1912) argued that abstract space is derived from the world of separate "objects" gradually identified by our perception. It is an elementary process, for perception must partition the continuous field which surrounds the body into objects upon which the body can *act*—to throw a "rock," to hoist a "bottle of beer," to grasp a "cube" which is "rotating." This fundamental perceptual partition into "objects" and "motions" is reified and extended in thought. The separate "objects" in the field are refined to the notion of the continuum of points or positions. As an object moves across this continuum, as for example, my hand moving across the desk from point A to point B, it is conceived to describe a *trajectory*—a line—consisting of the points or positions it traverses. Each point momentarily occupied is conceived to correspond to an "instant" of time. Thus arises the notion of abstract time—the series of instants—itself simply another dimension of the abstract space. This space, argued Bergson, is in essence a "principle of infinite divisibility." Having convinced ourselves that this motion is adequately described by the line/trajectory the object traversed, we can break up the line (space) into as many points as we please. But the concept of motion this implies is inherently an infinite regress. To account for the motion, we must, between each pair of points supposedly occupied by the object, re-introduce the motion, hence a new (smaller) trajectory of static points—ad infinitum. It is the stuff of Zeno and his paradoxes, yet this framework has lurked ubiquitously beneath our notion of the perception of form.

Lynds (*Foundations of Physics Letters*, 2003) has made similar arguments to Bergson's, again arguing that there is no precise static instant in time underlying a dynamical physical process. If there were such, motion and variation in all physical magnitudes would not be possible, as they would be frozen static at that precise instant, and remain that way. In effect, such an instant would imply a momentarily static universe. Such a universe is incapable of change, for the universe itself could not change to assume another static instant. Consequently, at no time is the position of a body (or edge, vertex, feature, etc.) or a physical magnitude precisely determined in an interval, no matter how small, as at no time is it not constantly changing and undetermined. It is by this very fact—that there is not a precise static instant of time underlying a dynamical physical process or motion—that variation in magnitudes is possible; it is a necessary tradeoff—precisely determined values for continuity through time. It is only the human observer, Lynds notes, who imposes a precise instant in time upon a physical process. Thus, there is no equation of physics, no wave equation, no equation of motion, no matter how complex, that is not subject to this indeterminancy.[10]

[10] Lynds, his reviewers and consultants (e.g., J.J.C. Smart) are apparently unaware of his total precedence by Bergson. However, Lynds' position on, (a) relativity and (b) his acceptance of the concept that the motion of time is merely illusion produced by our senses would both be rejected by Bergson. I have not discussed the static, 4-D block conception of space–time here for reasons of space, but I note simply that Rakić (1997) has proven that Minkowski space–time lacks ontological status. A temporal theory with no ontological status is not relevant to psychology.

With this view, there can be no static form at any instant, precisely because this static instant does not exist. The brain cannot base its computations on something that, to it, does not exist. The brain is equally embedded in the transforming matter-field, i.e., it is equally a part of this indeterminancy. It can only be responding to invariance over change.

Approaches such as Ullman's, though continuous in theory, vectored form to the instant, *t*, in which the model, *M*, resides. This instant is infinitely thin, in fact, infinitely divisible. To account for any form of a time-extended event, these infinitely thin instants must be stored successively in some "memory." Equivalently, the brain is conceived to be taking samples or snapshots of the rotating cube. A sample is taken, the computations and the feature analysis (into vertices, lines, geons, etc.) to construct a perceived form begin, then another sample is taken, etc. The samples are stored in memory. Essentially, as Turvey (1977) noted, the effect (in either case) is to create a trajectory—a series of points (snapshots) laid out on a line, i.e., in a static space, like snapshots laid out upon a desktop. It should be noted that this creates an instant practical problem. How, Turvey asked, could the sampling rate of the sampler be pre-adjusted to the symmetry period of the rotating cube? The sampler would have to know beforehand the revolution rate of any cube it possibly encountered so it could take samples at an integral multiple of the symmetry period! But to make matters more impossible, what if there were two cubes rotating at different rates, or three?

The energy model, with its spatiotemporal filters, provides a dynamic embedding of the computation of form. As the form of the velocity field defined over the figure is computed, instant by instant, the impression is given that this supports the continuous change of the form, e.g., the twisting non-rigidity of the rotating ellipsoid or the downwards motion of the trapezoid. But for this impression to hold, we, as consumers or students of the model, must create an *implicit* buildup of samples (instants) over time. The initial computation or value of the velocity field of the rotating, non-rigid ellipsoid is long in the past as the next one is computed, then the next, etc. For us to make any sense of the output of the algorithm *as an experience*, the samples are implicitly being stacked (let us say, subconsciously) along a fourth dimension or in some provisional memory store. There is no accounting in the theory for the primary memory that would support the experience of the twisting ellipsoid. As noted earlier, the continuity of our own consciousness is being presumed.

Formulating the above question, with Domini and Caudek (2003), as simply one of integration across space and time, fails to remove the problem of the memory implied. But there is another question also: what is the time scale at which the samples are being stacked? Are they successive instants of a "buzzing" fly, the successive (smaller) instants of a heron-like fly, the successive (yet smaller) instants of the crystalline vibrations of a "molecular" fly, etc? In this context, it can be seen that if we frame the problem of time-extended form as simply one of integration, we fail to account for the concrete imposition of the scale over which the integration is effected.

3. Consciousness and time-extended form

The purpose of the samples stored in a memory is to account for motion, i.e., to account for the time-extended perception of *events*—the cube as "rotating" or the fly, with his wings oscillating, as "buzzing" by. The rotating cube is inherently a time-extended experience. We have seen the information defining it must have time-extent. A sampling process, whether implicit or explicit, has serious difficulties, not only practically for the perception of form, but for the very fact that a series of static points a motion does not make—we can never reconstruct a motion from a set of immobilities. Should we manage to construct a vast composite of samples in the 3-D space of the brain, either of the rotating cube or of the wing motions of the buzzing fly, "who" now looks at this static set? How does it become a motion? Do we posit some internal scanner? If so, how does the scanner register motion? The regress begins again. It is the temporal form of the homunculus regress.

The brain, with its memory, is considered the safe place to store these samples, to preserve them from the non-existence of the past, for the brain, as *matter*, is always "present." What is the time-extent of this present? Is it the life span of some sub-particle, e.g., 10^{-9} ns, or even less? Is it 1 s? We are storing the samples in the brain because it is matter. We are quite confident of this, yet our understanding of matter has never been settled. Were we to apply the logic of this philosophy to the matter-field and its transformation in time, we would have the following: We imagine the whole of universal space as a 3-D "Cube" having an infinitesimal extent in time. The Cube, existing only an "instant" of our abstract time, instantly ceases to exist. Another is instantaneously generated, becomes non-existent, then another, etc. To construct a 4-D Cube of sufficient extent along a 4th dimension to account for the perception of the event of the "rotating" cube or of the "buzzing" fly, we would have to invoke some "force," glue or "memory" to stick an immensely long string of the instantaneous Cubes together. But having invoked this "force," how do we now limit the extent of the 4-D Cube, how do we keep it pared down to size? Is it infinite in extent? If not, by what principle is this extent limited? The entire lifetime of the fly is not seen, just a certain limited extent of its buzzing. The entire rotational history of the cube is not perceived, but clearly enough of a 4-D extent to specify its form.

One cannot simply appeal then to "memory" to solve the problem of the time-extent of consciousness. One cannot simply appeal to "continuous processes." This is routinely done. Consider Taylor (2002):

> The features of an object, bound by various mechanisms to activity in working memory, thereby provide the content of consciousness of the associated object... In these [neural activity loops], neural activity "relaxes" to a temporally stable state, therefore providing the extended temporal duration of activity necessary for consciousness... (Taylor, 2002, p. 11)

Assumed here again is the continuity of our own consciousness to explain the time-extent of consciousness. Again, what is the 4-D or time-extent of any of this neural activity?

We are misled by the computer metaphor. The machine takes samples of the external world; it maintains them in its memory; its continuous processes are conceived as executing programs over time. But what is the 4-D extent of the machine? What, in fact, is the 4-D extent of the Zero Point Field in which the machine is embedded? Again, 10^{-9} ns? One second? Infinite? At the best, again, we have lent to these processes of the machine only the continuity and extension in time of our own consciousness. This approach does not account for primary memory—the time-extension of perceived events—buzzing flies, rotating cubes, mellow-sounding violins.

3.1. Non-differentiable time-motion

Ultimately then, the problem of time-extended primary perception involves a metaphysical problem, i.e., a problem in our notions of space, time and motion. This should not surprise us.

Physics has been struggling with the same root problem. The concept of the "trajectory" of a motion has long since been abandoned. As De Broglie (1947/1969) noted, the essence of Heisenberg's uncertainty is that the projection of a motion to a position/point (or to a series of points on a line) in the abstract continuum of positions merely results in immobility—we have lost the motion. Nottale (1996), noting Feynman and Hibbs (1965) demonstration that the motion of a particle is continuous, but not differentiable, now questions the hitherto fundamental assumption of the differentiability of space–time. To say this in another way, the global evolution of the matter-field over time is seen as non-differentiable; it cannot be treated as an infinitely divisible series of states. The attempt to project the motion of the cube via samples or snapshots to a series of static positions (and static features) relies on just this differentiability assumption.

Bergson (1896/1912) visualized this non-differentiable motion of the matter-field in terms of a melody, the flowing notes of which interpenetrate, each note being a reflection of the preceding series, an organic continuity. Referring to the relativistic end-result of abstract space and time, he would argue:

> Of what object, externally perceived, can it be said that it moves, of what other that it remains motionless? To put such a question is to admit the discontinuity established by common sense between objects independent of each other, having each its individuality, comparable to kinds of persons, is a valid distinction. For on the contrary hypothesis, the question would no longer be how are produced in given parts of matter changes of position, but how is effected in the whole a change of aspect... (1896/1912, p. 259).

The "motions" of "objects" now become changes or *transferences of state* in this melodic, global transformation of the whole. From this perspective, "primary memory" becomes a property of the field itself and of its melodic motion.

The brain is embedded within this field and its non-differentiable motion. Given this, and as Gibson (1966) can be read to imply (Robbins, 2000), we can view the dynamical pattern it supports as "specific" to an indivisible, variable (scaled), time-extent of this field—a "buzzing" fly, or a heron-like fly, a spinning 16-edged cylinder-cube, or a static cube, or according to the invariance laws constraining this dynamical pattern, a non-rigid, wobbly cube.

From this perspective, and given the discussion of scale and invariance, certain general constraints on the computational system or device required to support the *experience* of dynamic form can be proposed as follows:

- The global dynamics of the system must be proportionally related to the events of the matter-field such that a time-scale is defined upon this field.
- The operative dynamics of the system must be structurally related to the events of the field, i.e., reflective of the invariance laws defined over the time-extended events of the field.
- The operative dynamics of the system must be an integral part of the indivisible motion of the field in which it is embedded.

I have used "operative" here. Expositors of the computer model (Dietrich & Markman, 2000; Prinz & Barsalou, 2000) have argued that since the syntactic, discrete operations of their programs (the effective, operative aspect of the system) rest upon the *real* dynamics of the machine, this is sufficient. This is not the case. It is not possible for this effective syntax, being merely the manipulation of abstract objects in an abstract space, to embody these constraints (Robbins, 2002) and therefore to support the specification of form.

These constraints, by themselves, leave us with a critical question: how does this dynamics—this "specification"—even under these constraints, result in the external *image* of form—the *image* of the rotating cube or the buzzing fly? This *external* image *is* our very experience of form. The "specification" itself, albeit embodying computations of velocity fields under probabilistic constraints, is still simply neural patterns transforming, chemical flows, spatiotemporal filters firing. How can we conceive of the end result of this specification—the *image* of a dynamic form—as equally as physical a result as the neural flows supporting it? All depends on how we conceive of this specification. For the sake of a degree of closure, I offer briefly a conjecture developed in a bit more detail elsewhere (Robbins, 2002) that is a consequence, I believe, of the foregoing considerations on non-differentiable time, primary memory and dynamic form.

3.2. The specification of form as image

Let us note that the "specification" supported by this dynamics is to the *past*, i.e., to past "states" of the transforming matter-field. The external events the brain is processing—the wing-beats of the fly, the motion-cycles of the cube, and all the micro-events comprising these motions—have long since come and gone. Yet the relativity principle we have discussed in the context of Messrs. A and B, the non-differentiable or melodic motion of the field, the fictional status of present "instants" that cease instantly to exist—all tell us that this past-specification is possible.

Bergson, as did Bohm (1980), saw his globally transforming field as having holographic properties. It is, in essence, a vast, dynamic interference pattern (see also Beckenstein, 2003). Assuming this, the dynamical, resonant pattern supported by the brain can then be conceived as supporting a modulated *reconstructive wave* passing thru this holographic field.[11]

A reconstructive wave of a certain frequency, passing through a hologram plate (i.e., through the interference pattern stored on this plate), specifies a virtual image of an object (a coffee cup for example) in 3-D space (Fig. 14). The interference pattern stored on the plate is the record of two waves—the original object wave reflected from the cup, and a reference wave of a precise or coherent frequency (usually emitted by a laser). By modulating the reconstructive wave to a second frequency, a different stored object (e.g., a cube) can be specified. An imprecise reconstructive wave, containing both frequencies (i.e., a wave of lesser coherence), will specify a composite image of cube and cup. So too, the brain, seen as a modulated reconstructive wave "passing through" the matter-field, would continuously specify a changing virtual image of the past motion of this field. Due to the hierarchical dynamics of the brain, the dynamical pattern (or attractor) is in a certain *proportionality* relative to the micro-events of the matter-field, and therefore the image is specific to a scale of time—a spinning, fuzzy cylinder-cube, or a barely rotating cube. Assuming

[11] Conceiving of the brain as a wave is certainly not unprecedented. Yasue, Jibu, and Pribram (1991), see the evolving brain states in terms of complex valued wave flows, where constraints on the brain's (state) evolution are elegantly represented by Fourier coefficients of the wave spectrum of this formulation. Glassman (1999), for example, attempts to account for the limited capacity of working memory by viewing the brain, globally, as a set of waves whose frequencies are confined to a single octave.

YCCOG 705
5 August 2004 Disk Used
ARTICLE IN PRESS
No. of Pages 27, DTD = 5.0.1
Jaya (CE) / Selvi (TE)

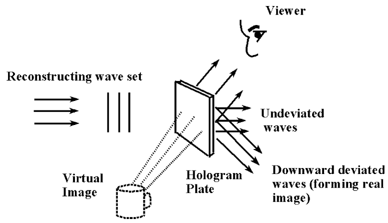

Fig. 14. Holographic reconstruction. A set of plane waves of the same frequency, f_1, as the original reference wave used to store the interference pattern (hologram) strikes the plate and is diffracted in different directions. The upward rising wave set specifies the virtual image of the (stored) object. Another reconstructive wave, modulated to a different frequency, f_2, can reconstruct a different stored wave front, perhaps the image of a cube, etc. (cf. Kock, 1969).

that the resonant, dynamic pattern supporting this wave is constrained by fundamental invariance laws, the image takes a structure defined by these laws.[12]

This would be a concrete realization of Gibson's abstract "direct specification" of events or of dynamic forms. It is a direct realism that is not simply a naïve realism. The image is always an optimal function of the invariance information available in the field in conjunction with invariance laws (constraints) built into the brain's design. It is a specification of the past motion of the field given the best available information within the field and given the intrinsic uncertainty of "measuring" this field due to its temporal motion. Being a specification of the past, it is always, already a memory, a memory based in the primary memory supported by the non-differentiable time-evolution of the matter-field itself.

4. Conclusion

We have journeyed from a meditation on the difficulties of explaining the perception of dynamic form, to the abstract time underlying our thought on the subject, to the non-differentiable time towards which physics is pointing, and to a conjecture, apparently within the implications of such a melodic time, on a possible method of specification of the changing, time-scale specific, "external" image of forms within the matter-field. Given that the brain and its processes are embedded in a global, non-differentiable motion of the matter-field, then it is within the implications of this conception of time and motion that we may have to look for the explanation of something as simple as the time-extended perception of a rotating cube.

[12] There is no homunculus here, no viewer of the image as in Fig. 14. As there are not separate "objects" in the field, the distinction between subject and object cannot be in terms of space. Rather, as Bergson argued, the distinction is only in terms of *time*. As the dynamics places successive scales on time, where the external "fly" moves from waves in the field undifferentiated from an observer, to an ensemble of vibrating molecules, to a motionless form, to a heron-like fly barely moving his wings, to the buzzing being of normal scale, object differentiates from subject. As there is an elementary awareness implicit in the holographic properties of the field at the null scale of time, the specification is to a past, time-scaled form and subset of this elementary awareness.

5. Uncited references

Shaw, McIntyre, and Mace (1974), Watson and Ahumada (1985).

References

Adelson, E., & Bergen, J. (1985). Spatiotemporal energy model of the perception of motion. *Journal of the Optical Society of America, 2*(2), 284–299.

Ames, A. (1951). Visual perception and the rotating trapezoidal window. *Psychological Monographs*, 65 (7, Whole No. 234).

Baddeley, A. (1986). *Working memory*. Oxford: Clarendon Press.

Beckenstein, J. (2003). Information in the holographic universe. *Scientific American, 289*(2), 58–66.

Beghi, L., Xausa, E., & Zanforlin, M. (1991). Analytic determination of the depth effect in stereokinetic phenomena without a rigidity assumption. *Biological Cybernetics, 65*, 425–432.

Beghi, L., Xausa, E., De Biasio, C., & Zanforlin, M. (1991). Quantitative determination of the three-dimensional appearances of a rotating ellipse without a rigidity assumption. *Biological Cybernetics, 65*, 433–440.

Bergson, H., (1896/1912). *Matter and Memory*. New York: Macmillan.

Bohm, D. (1980). *Wholeness and the implicate order*. London: Routledge and Kegan-Paul.

Braunstein, M. L. (1976). *Depth perception through motion*. New York: Academic Press.

Carlton, E., & Shepard, R. (1990a). Psychologically simple motions as geodesic paths I. Asymmetric objects. *Journal of Mathematical Psychology, 34*, 127–188.

Carlton, E., & Shepard, R. (1990b). Psychologically simple motions as geodesic paths II. Symmetric objects. *Journal of Mathematical Psychology, 34*, 189–228.

De Broglie, L., (1947/1969). The concepts of contemporary physics and Bergson's ideas on time and motion. In: P.A.Y. Gunter (Ed.), *Bergson and the evolution of physics*, University of Tennessee Press.

Dietrich, E., & Markman, A. B. (2000). Cognitive dynamics: computation and representation regained. In E. Dietrich & A. B. Markman (Eds.), *Cognitive dynamics: Conceptual and representational change in humans and machines*. New Jersey: Erlbaum.

DeValois, R. L., Albrecht, D. G., & Thorell, L. G. (1978). Cortical cells: bar and edge detectors, or spatial frequency filters?. In S. J. Cool & E. L. Smith (Eds.), *Frontiers of visual science*. New York: Springer-Verlag.

Domini, F., Vuong, Q. C., & Caudek, C. (2002). Temporal integration in structure from motion. *Journal of Experimental Psychology: Human Perception and Performance, 28*(4), 816–838.

Domini, F., & Caudek, C. (2003). 3-D structure perceived form dynamic information: A new theory. *Trends in Cognitive Sciences, 7*, 444–449.

Domini, F., & Braunstein, M. L. (1998). Recovery of 3D structure from motion is neither Euclidean or affine. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1273–1295.

Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., & Reitboeck, H. J. (1988). Coherent oscillations: A mechanism of feature linking in the visual cortex? Multiple electrode and correlation analyses in the cat. *Biological Cybernetics, 60*, 121–130.

Epstein, W. (1995). The metatheoretical context. In W. Epstein & S. Rogers (Eds.), *The perception of space and motion*. New York: Academic Press.

Feynman, R. P., & Hibbs, A. R. (1965). *Quantum mechanics and path integrals*. New York: MacGraw-Hill.

Fischer, R. (1966). Biological time. In J. T. Fraser (Ed.), *The voices of time*. New York: Braziller.

Freyd, J. J. (1987). Dynamic mental representations. *Psychological Review, 94*, 427–438.

Gibson, J. J. (1950). *The perception of the visual world*. Boston: Houghton-Mifflin.

Gibson, J. J. (1966). *The senses considered as visual systems*. Boston: Houghton-Mifflin.

Gibson, J. J. (1975). Events are perceived but time it not. In J. T. Fraser & N. Laurence (Eds.), *The study of time II*. New York: Springer-Verlag.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton-Mifflin.

Glassman, R. B. (1999). Hypothesized neural dynamics of working memory: Several chunks might be marked simultaneously by harmonic frequencies within an octave bank of brain waves. *Brain Research Bulletin, 30*(2), 77–93.

Hildreth, E. C. (1983). *The Measurement of Visual Motion*. Cambridge, MA: MIT Press.

Hildreth, E. C., Ando, H., Andersen, R. A., & Treue, S. (1995). Recovering three-dimensional structure from motion with surface reconstruction. *Vision Research, 35*, 117–137.

Hoaglund, H. (1966). Some bio-chemical considerations of time. In J. T. Fraser (Ed.), *The voices of time*. New York: Braziller.

Hochberg, J. (1998). Gestalt theory and its legacy. In J. Hochberg (Ed.), *Perception and cognition at century's end*. New York: Academic Press.

Hoffman, D. D. (1998). *Visual intelligence*. New York: Norton.

Hoffman, D. D., & Flinchbaugh, B. E. (1982). The interpretation of biological motion. *Biological Cybernetics, 42*, 195–204.

Hubel, D., & Wiesel, T. N. (1959). Receptive fields of single neurons in the cat's striate cortex. *Journal of Physiology, 148*, 574–591.

Hubel, D., & Wiesel, T. N. (1978). Brain mechanisms in vision. *Scientific American, 241*, 150–162.

Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Reviews, 12*, 487–519.

Ittleson, W. (1962). Perception and transactional psychology. In S. Kock (Ed.), *Psychology: A study of science*. New York: McGraw-Hill.

James, W. (1890). *Principles of psychology*. New York: Holt and Co.

Kock, W. E. (1969). *Lasers and holography*. New York: Doubleday-Anchor.

Liu, Z. (2003). On the principle of minimal relative motion—the bar, the circle with a dot, and the ellipse. *Journal of Vision, 3*, 625–629.

Lynds, P. (2003). Time and classical and quantum mechanics: Indeterminacy versus discontinuity. *Foundations of Physics Letters, 16*(4), 343–355.

Mussati, C. L. (1924). Sui fenomeni stereocinetici. *Archivo Italiano di Psycologia, 3*, 105–120.

Nakayama, K. (1998). Vision fin de siècle: A reductionistic explanation of perception for the 21st century?. In J. Hochberg (Ed.), *Perception and cognition at century's end*. New York: Academic Press.

Nottale, L. (1996). Scale relativity and fractal space–time: applications to quantum physics, cosmology and chaotic systems. *Chaos, Solitons and Fractals, 7*, 877–938.

Pittenger, J. B., & Shaw, R. E. (1975). Aging faces as viscal elastic events: Implications for a theory of non rigid shape perception. *Journal of Experimental Psychology: Human Perception and Performance, 1*, 374–382.

Pomerantz, J. R. (1983). The rubber pencil illusion. *Perception and Psychophysics, 33*, 365–368.

Prinz, J., & Barsalou, L. (2000). Steering a course for embodied representation. In E. Dietrich & A. B. Markman (Eds.), *Cognitive dynamics: Conceptual and representational change in humans and machines*. New Jersey: Erlbaum.

Rakić, N. (1997). Past, present, future, and special relativity. *British Journal for the Philosophy of Science, 48*, 257–280.

Reichardt, W. (1959). Autocorrelation and the central nervous system. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 303–318). Cambridge: MIT Press.

Robbins, S. E. (2000). Bergson, perception and Gibson. *Journal of Consciousness Studies, 7*(5), 23–45.

Robbins, S. E. (2002). Semantics, experience and time. *Cognitive Systems Research, 3*(3), 301–337.

Rowe, M. H., & Stone, J. (1980). Parametric and feature extraction analyses of the receptive fields of visual neurons. *Brain, Behavior and Evolution, 17*, 103–122.

Shaw, R. E., McIntyre, M., & Mace, W., (1974). The role of symmetry in event perception. In Robert B. MacLeod & Herbert L. Pick (Eds.), *Perception: Essays in Honor of James J. Gibson* (pp. 276–310).

Shaw, R. E., & McIntyre, M. (1974). The algoristic foundations of cognitive psychology. In D. Palermo & W. Weimer (Eds.), *Cognition and the symbolic processes*. New Jersey: Lawrence Erlbaum Associates.

Shaw, R. E., & Mace, W. (in press). The value of oriented geometry for ecological psychology.

Singer, W. (1998). Consciousness and the structure of neural representations. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 353*, 1829–1840.

Taylor, J. G. (2002). From matter to mind. *Journal of Consciousness Studies, 9*(4), 3–22.

YCCOG 705
5 August 2004 Disk Used
ARTICLE IN PRESS
No. of Pages 27, DTD = 5.0.1
Jaya (CE) / Selvi (TE)

Todd, J. F. (1982). Visual information about rigid and non-rigid motion: A geometric analysis. *Journal of Experimental Psychology: Human Perception and Performance, 8*, 238–251.

Todd, J. F. (1995). Visual perception of 3-D structure. In W. Epstein & S. Rogers (Eds.), *The perception of space and motion*. New York: Academic Press.

Turvey, M. (1977). Contrasting orientations to the theory of visual information processing. *Psychological Review, 84*, 67–88.

Ullman, S. (1979a). *The interpretation of visual motion*. Cambridge: MIT Press.

Ullman, S. (1979b). The interpretation of structure from motion. *Proceedings of the Royal Society of London Series B, 203*, 405–426.

Ullman, S. (1984). Maximizing rigidity: The incremental recovery of 3-D structure from rigid and non-rigid motion. *Perception, 13*, 255–274.

Ullman, S. (1986). Competence, performance and the rigidity assumption. *Perception, 15*, 644–646.

Von Hornbostel, E. M. (1922). Über optische inversion. *Psychologische Forschung, 1*, 130–156.

Wallach, J., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology, 45*, 205–217.

Wallach, J., Weiss, H., & Adams, P. A. (1956). Circles and derived figures from rotation. *American Journal of Psychology, 9*, 48–59.

Watson, A. B., & Ahumada, A. J., (1983). A look at motion in the frequency doman. NASA Tech memo, TM 84352.

Watson, A. B., & Ahumada, A. J. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A, 2*, 322–341.

Webb, J. A., & Aggarwal, J. K. (1981). Visually interpreting the motions of objects in space. *Computer, 8*, 40–46.

Weiss, Y., & Adelson, E., (1998). Slow and smooth: A Bayesian theory for the combination of local motion signals in human vision. MIT A.I. Memo No. 1624.

Weiss, Y., Simoncelli, E., & Adelson, E. (2002). Motion illusions as optimal percepts. *Nature Neuroscience, 5*, 598–604.

Yasue, K., Jibu, M., & Pribram, K. H. (1991). A theory of non-local cortical processing in the brain. In K. H. Pribram (Ed.), *Brain and perception*. New Jersey: Erlbaum.