Challenges of vision theory: Self-organization of neural mechanisms for stable steering of object-grouping data in visual motion perception

Jonathan A. Marshall
Department of Computer Science, CB 3175, Sitterson Hall
University of North Carolina, Chapel Hill, NC 27599-3175, U.S.A.
919-962-1887, marshall@cs.unc.edu

Abstract.

Psychophysical studies on motion perception suggest that human visual systems perform certain *nonlocal* operations. In some cases, data about one part of an image can influence the processing or perception of data about another part of the image, across a *long spatial range*. In others, data about *nearby* parts of an image can *fail* to influence one another strongly, despite their proximity. Several types of nonlocal interaction may underlie cortical processing for accurate, stable perception of visual motion, depth, and form:

- *trajectory-specific* propagation of computed moving stimulus information to successive image locations where a stimulus is predicted to appear;
- *grouping* operations (establishing linkages among perceptually related data);
- *scission* operations (breaking linkages between unrelated data); and
- *steering* operations, whereby visible portions of a visual group or object can control the representations of invisible or occluded portions of the same group.

Nonlocal interactions like these could be mediated by long-range excitatory horizontal intrinsic connections (LEHICs), discovered in visual cortex of several animal species. LEHICs often span great distances across cortical image space. Typically, they have been found to interconnect regions of *like specificity* with regard to certain receptive field attributes, e.g., stimulus orientation.

It has recently been shown that several visual processing mechanisms can self-organize in model recurrent neural networks using unsupervised "EXIN" (excitatory+inhibitory) learning rules. Because the same rules are used in each case, EXIN networks provide a means to unify explanations of how different visual processing modules acquire their structure and function. EXIN networks learn to *multiplex* (or represent simultaneously) multiple spatially overlapping components of complex scenes, in a context-sensitive fashion. Modeled LEHICs have been used together with the EXIN learning rules to show how visual experience can shape neural mechanisms for nonlocal, context-sensitive processing of visual motion data.

Nonlocal Interactions Mediate Perceptual Stability.

When we see one visual object move behind another, becoming partially occluded, the shape of its projection on our retina is drastically altered. Yet we generally perceive the object as remaining whole and undistorted, even though we cannot see all of it. How can our visual systems allow us to perceive an object, becoming occluded and unoccluded, as remaining the same? How can our perception of objects remain stable, even though we see scenes containing many objects at once whose images interact, producing radical distortions? These issues of *perceptual stability* are central to vision theory.

These challenging questions can be faced by investigating the role of *nonlocal interactions* in specific aspects of visual perception. By "nonlocal" interactions I refer to two sorts of processing:

- *Type A*: ways in which data about one part of an image can influence the processing or perception of data about another part of the image, across a *long spatial range*. Example: in Kanizsa

Invited paper, in *Stochastic and Neural Methods in Signal Processing, Image Processing, and Computer Vision*, Su-Shing Chen, Ed., Proceedings of the SPIE 1569, San Diego, July 1991, pp.200–215.

subjective triangles (Kanizsa, 1979), subjective contours are perceived across regions lacking objective contrast. The colinear alignment of certain edges is detected despite large spatial gaps.

- *Type B*: ways in which data about *nearby* parts of an image can *fail* to influence one another strongly, despite their proximity. Example: when one visual object occludes another, the perceived visual properties of their mutual image boundary (such as its motion) become closely linked to those of the occlud*ing* object and become entirely unlinked from those of the occlud*ed* object (Nakayama, Shimojo, & Silverman, 1989). The boundary becomes anisotropic, functionally "attached" to one side and disengaged from the other side.

Nonlocal Phenomena in Visual Perception.

A surprisingly diverse array of perceptual phenomena exhibit nonlocal interactions and appear to require nonlocal processing mechanisms.

Percept #1: Kinetic Subjective Contours.

Kellman and Cohen (1984) generated moving subjective contours by displaying a rotating white triangle that intermittently occludes parts of black circles on a white background (Figure 1). Because there is no visible contrast between the white parts of the background and the white triangle, no objective contours can be seen there. Objectively, the black parts of the background have moving white nicks or chunks taken out of them. But Kellman and Cohen found that observers perceive the figure clearly and continuously as a coherent whole triangle.

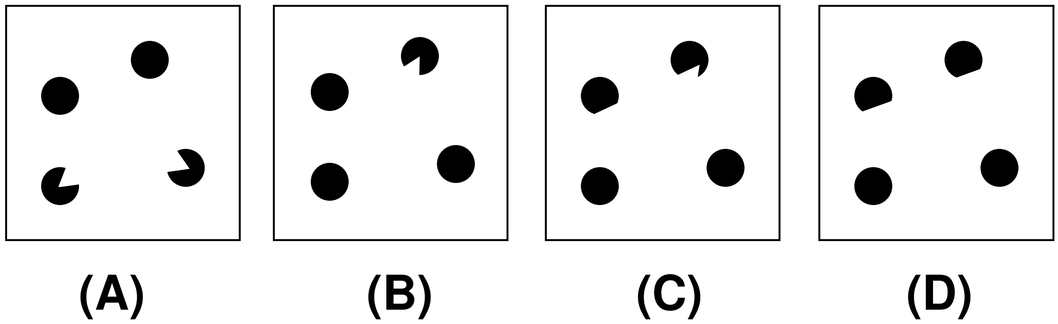

Figure 1: Kinetic subjective contours. Four frames of a rotating subjective triangle display.

These coherent percepts are obtained even when the display is arranged so that only one corner of the triangle is visible at a time. The subjective boundaries of the moving triangle are still clearly perceived.

This phenomenon argues against a straightforward colinear-completion scheme like that of Grossberg and Mingolla (1985ab) or von der Heydt, Peterhans, and Baumgartner (1984), because the colinearity-detectors of such schemes require two colinear edges to be simultaneously present in order to generate a bridging subjective contour. Thus, simple line-completion processes cannot explain the phenomenon of kinetic subjective contours: two parts of an edge to be completed might not appear simultaneously, and thus their colinear alignment might not even be detected at any given moment. Moreover, because the two edges to be bridged may appear at different times, they may never even be aligned spatially.

What causes a subjective corner or contour to continue moving perceptually even when it is objectively invisible? How do the invisible portions of a figure stay "on track," contributing to the percept of the whole figure, rather than flying off in the wrong direction? The motion of the representation of the invisible portions of a figure seems to be steered or controlled by the motion of the visible portions.

The representation of the invisible portions of a figure seems remarkably well-regulated: invisible features continue to be grouped along with visible features into a representation of a whole *object*, whose parts bear an internally coherent relation to one another. Some researchers (e.g., Parks, 1986; Prazdny, 1986) have suggested that figural information, such as the presence of occluding surfaces, may be more central than edge information in the perception of subjective contours.

The representation of each object is functionally segregated from that of other objects; motion signals generated by visible portions of the triangle apply only to the triangle, and not to the background circles.

These properties suggest that edge completion and figural completion processes in human vision operate using a dynamic framework in which information picked up about a figure at one location can be "carried" along with the motion of the figure to other locations.

Percept #2: Motion Viewed Through a Slit (Parks' Camel).

Parks (1965) devised a display in which a moving figure (e.g., a drawing of a camel) was viewed through a narrow slit in a screen (Figure 2) for 250–500 ms. Even though only a small slice of the figure was visible at any single moment, observers are able to see the figure as a whole briefly after it has moved past the slit.

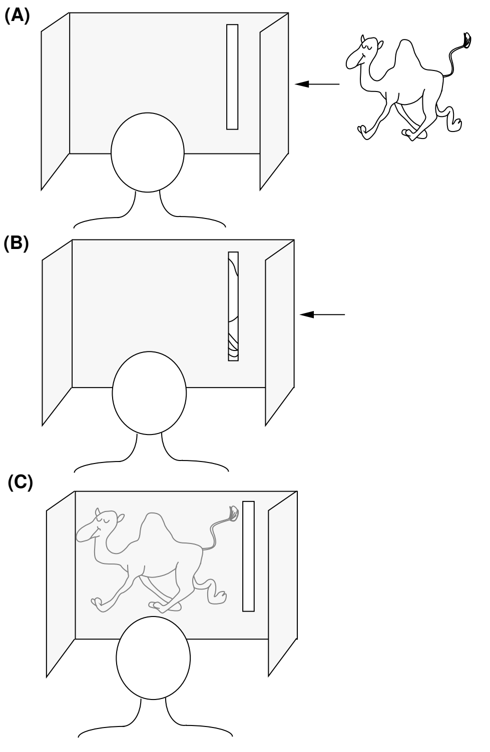

Figure 2: Parks' Camel. An image, such as a drawing of a camel (A), is viewed as it moves behind a narrow slit (B). Under proper conditions, observers report seeing the figure as a whole (C), possibly with some elongation (dotted figure of camel).

It seems that our visual systems can arrange the narrow "slices" visible at successive instants into the proper order so that the image of the entire figure is reconstructed and seen. This result suggests that the new incoming motion signals from the narrow slit cause the representation of previously-recorded slices to "move over" so that the slices will be correctly lined up to accommodate the new information. In other words, the motion of the visible portions of an object drives the motion of the representation of the invisible portions.

Parks ruled out explanations based on "retinal painting" and on eye movements by moving two figures simultaneously in opposite directions behind the same slit; both could be seen afterward.

Shimojo, Silverman, and Nakayama, (1988) found that *binocular* viewing of moving figures through a slit lets observers determine the depth of objects. Visual systems thus use the timing relationships between corresponding slices from the left- and right-eye images in computing binocular disparity.

Percept #3: Visual Inertia.

Motion signals generated between two successive frames of a movie-like display bias the motion perceived between the next frames (FIGURE 3) (Anstis & Ramachandran, 1987). A visual stimulus appears to have *inertia*; its motion at one instant generates an expectation, or bias, or priming signal, for it to continue to move along the same trajectory.

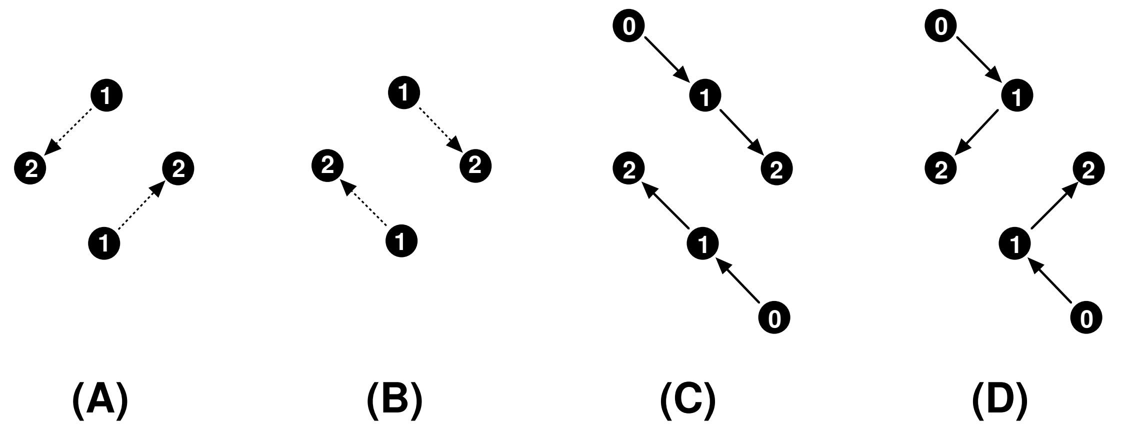

Figure 3: Visual Inertia. Dots are flashed on a display screen at the times indicated by numbers. When alternating horizontal/vertical pairs of dots are flashed at the vertices of a diamond shape, observers report a bistable motion percept, either (A) or (B), at any moment. But when the same dot flashes are preceded by priming dots (labeled 0), observers report a monostable motion percept, always (C), never (D).

Two dots can generate priming effects in opposite directions of motion; therefore, the priming cannot be a global effect. Rather, each stimulus specifically biases its own motion, and not that of other stimuli. The bias propagates only in the same specific direction, speed, and locations as the stimulus's trajectory.

Percept #4: Representation of Invisible Features.

In a study of "illusory occlusion," Ramachandran, Inada, & Kiama (1986) flash an array of random dots, and then flash the entire array again, except for one removed dot, in the same configuration, but at a slightly shifted location. The position that the removed dot might have occupied in the second frame is covered by an opaque piece of cardboard (FIGURE 4).

Even though the dot is visible only in one of the two frames, it is generally perceived to move in synchrony with the rest of the dots, appearing to become occluded as it moved "behind" the cardboard. Yet when the cardboard is removed, that dot appeared simply to flash; the percept of motion for that dot is greatly reduced or eliminated.

This result suggests that visual systems can *assign* or *impute* grouped motion directly to an objectively stationary stimulus, when permitted by occlusion relations. Thus, under appropriate conditions, visual systems seem to prefer to treat disappearing stimuli as continuing to exist but becoming occluded, rather than ceasing to exist.

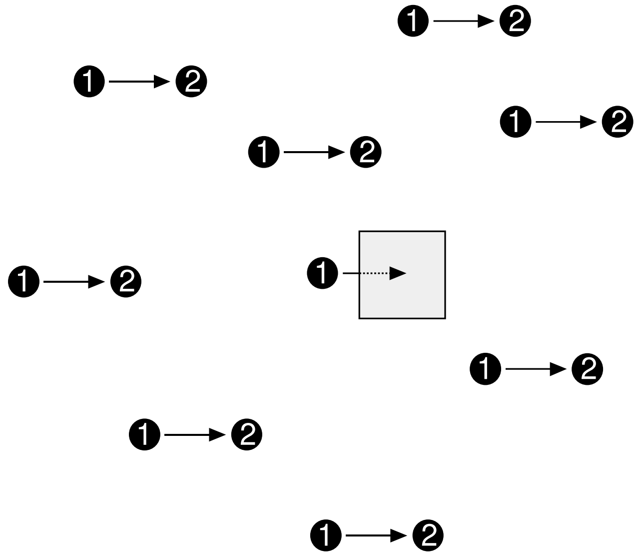

Figure 4: Illusory Occlusion. One dot's position at time=2 is occluded, yet the dot appears to move "behind" the occluder.

Percept #5: Dynamic Occlusion.

Accretion/deletion events give human observers clear percepts of relative depth. Kaplan (1969) created a random dot display in which motion of image regions caused the accretion and deletion of optical texture (dots) at the region boundaries (FIGURE 5). The region whose motion velocity is closest to the velocity of the boundary between the regions is seen as the occluding surface; the other region is seen as being occluded by it.

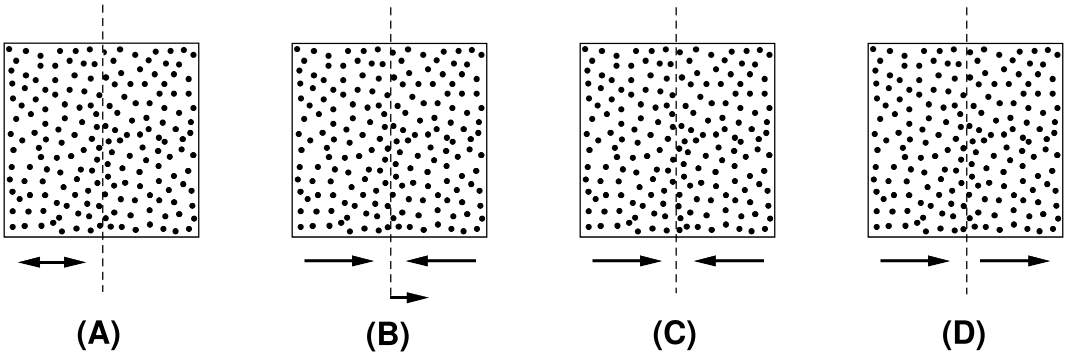

Figure 5: Dynamic Occlusion: accretion and deletion of optical texture. Randomly arranged dots are displayed. A vertical subjective boundary separates the display into two regions. (A) When the dots on the left side move back and forth together, disappearing and reappearing at the subjective boundary, the right side appears to occlude the left side. (B) When the subjective boundary moves in the same direction and speed as the dots on the left side, dots on the right accrete and are deleted, and the left side appears to occlude the right side. (C) When the dots on both sides move toward each other at the same speed, neither side appears to occlude the other. (D) When both sides move in the same direction and the same speed, and dots on the left are deleted as they reach the boundary, and new dots accrete on the right side, then the boundary itself appears to be a "thin string" or "thin pencil," lying in front of the moving dots.

A related effect can be noted in comparing a simple motion transparency display with a "bouncing dot" motion transparency display. In the simple motion transparency display (Ullman, 1979), a set of randomly arranged dots are shown, where half the dots are moving left and half right, at a single constant speed. Dots reaching the (invisible) left and right edges of the display disappear, and new moving dots appear along the left and right edges. In this case, observers report that the display appears to contain two transparently overlaid flat surfaces, one moving left and one right.

Ramachandran, Cobb, and Rogers-Ramachandran (1988) created a second display, identical to the first, except that dots simply reverse direction (or "bounce") at the left and right edges and retrace their paths, instead of being deleted and replaced. In this case, observers report a rotating-cylinder percept.

This phenomenon suggests that perceived 3-D structure (two sliding flat surfaces vs. one rotating cylinder) depends significantly on fine image details such as continuous existence vs. disappearance of dots at boundaries. These results suggest that accretion/deletion of individual visual features may be even more salient than optic flow *per se* in determining our percepts of 3-D structure.

Percept #6: Motion Transparency.

Random-dot rotating cylinder displays (Ullman, 1979) illustrate the capability of visual systems to allow different (even opposite) motion signal groupings to occur across the same image region. Visual processing mechanisms can thus implement *scission* operations, whereby surfaces can be functionally disengaged from one another so that motion signals in one do not interfere with those in the other.

Percept #7: Motion Sharpening Along Trajectories.

A moving visual stimulus appears smeared (with a comet-like tail) when it is viewed for a brief exposure (30 ms), yet perfectly sharp when viewed for a longer exposure (100 ms) (Burr & Ross, 1986; Van Essen & Anderson, 1987). Our ability to counteract smear at longer exposures suggests that our visual systems combine and sharpen motion information from multiple locations along a spatiotemporal trajectory that matches the motion of the stimulus (Barlow, 1979, 1981; Burr & Ross, 1986).

Analysis: Segmentation, Grouping, Continuity, and Communication.

Taken together, these psychophysical phenomena suggest the following hypotheses about key functional capabilities of visual systems.

Invisibility Representation Hypothesis. Visual systems can maintain a subjective representation of the objectively invisible portions of a figure (though perhaps only for a limited time). The invisible portions include occluded features, subjective contours, and "filled-in" surface data.

Visible/Invisible Grouping Hypothesis. Visual systems can group together as an internally coherent object the representations of the visible and invisible portions of a figure, to form a unitary object-centered representation that can be manipulated as a whole.

Coherence Hypothesis. Visual systems use coherence information, such as common fate, proximity, and similarity, in determining what features constitute each object. Features sharing similar characteristics are linked together. The linkages are maintained despite transformations such as shifts, rotations, and occlusions.

Invisibility Imputation Hypothesis. Visual systems interpolate or fill-in putatively missing information within each object representation, complete moving subjective contours even between colinearly aligned pairs of edges that are not simultaneously visible, and generate interpolated surface groupings from sparse texture.

Steering Hypothesis. The motion of the visible portions of an object "steers" or controls the motion of the representation of the object, including the invisible portions of the object.

Trajectory-Specific Propagation Hypothesis. Steering control is exerted in a trajectory-specific manner. Each object representation sends grouping and motion signals to successive object representations that it expects the object to activate. By providing prior state information about the object, these anticipatory, or predictive, signals help sharpen its future representation. The predictive signals also help buffer the transforming representation from erosion due to noise or to occlusion. In addition, they allow the visible portions of the object to steer the successive representations of the invisible portions. Because the predictive signals are transmitted along highly specific pathways, the predictive signals from different objects can operate with minimal mutual interference.

Scission Hypothesis. Object representations are segregated or isolated from one another, so that signals in one object grouping do not interfere with signals in another, even if the objects overlap spatially.

Nonlocal Structures in Visual Anatomy and Physiology.

Long-range excitatory horizontal intrinsic connections (LEHICs) have been identified in the visual cortex of a variety of animals (Blasdel, Lund, & Fitzpatrick, 1985; Callaway & Katz, 1990; Gabbott, Martin, & Whitteridge,, 1987; Gilbert & Wiesel, 1989; Luhmann, Martínez Millán, & Singer, 1986; Lund, 1987; Michalski, Gerstein, Czarkowska, & Tarnecki, 1983; Mitchison & Crick, 1982; Nelson & Frost, 1985; Rockland & Lund, 1982, 1983; Rockland, Lund, & Humphrey, 1982; Ts'o, Gilbert, & Wiesel, 1986).

The LEHICs exhibit several characteristics relevant to the hypotheses described above:

- They are highly specific, running anisotropically from one small patch of cortex to another.
- They typically run between cells with similar stimulus preferences.
- Some run for very long distances (e.g., 10 mm) across the cortex (Luhmann, Martínez Millán, & Singer, 1986).
- They are shaped adaptively through visual experience. "There is an inborn pattern of discrete horizontal connections in striate cortex which is shaped by visual experience and requires contour vision for its maintenance" (Luhmann, Martínez Millán, & Singer, 1986, p. 443). "The development of the horizontal connections in striate cortex occurs in at least two phases, an early phase during which connections are formed in excess and a later phase during which connections are again eliminated" (Ibid., p. 447). Callaway & Katz (1990) suggest that "pyramidal neurons making long-distance horizontal connections use activity cues to identify appropriate target cells having similar orientation selectivity" (p.1151).
- They may serve to predictively prime motion-sensitive cells: "This feedforward excitation by a neuron into whose receptive field an object has just entered could act as a facilitatory device to 'prime' (by increasing their response sensitivity) other neurons into whose receptive fields the object might eventually travel. These operations could also provide some predictive estimate of the future position of the object" (Gabbott, Martin, & Whitteridge, 1987, pp. 378–379).

Since LEHICs are found in area V1, similar connections are likely to be found within higher cortical levels as well, like areas MT and STS. LEHICs may allow linkages and transformations between groupings or between instances of a grouping to be learned and allow priming or control to be exercised distally. Possible mechanisms for these hypothesized functions of LEHICs are described below.

EXIN Networks – Powerful New Techniques for Nonlocal Prediction, Segmentation, and Grouping.

Self-Organizing Neural Networks for Visual Processing.

Self-organization methods, using unsupervised learning rules, have been useful in modeling the development of several aspects of low-level visual perception, including sensitivity to contrast (Linsker, 1986a), orientation (Bienenstock, Cooper, & Munro, 1982; Linsker, 1986b; Marshall, 1990f; von der Malsburg, 1973), motion (Coolen & Kuijk, 1989; Földiák, 1991; Marshall, 1989ab, 1990a; M.E. Sereno, 1986, 1987; M.E. Sereno, Kersten, & Anderson, 1988; M.I. Sereno, 1989; M.I. Sereno & M.E. Sereno, 1990), length (Marshall, 1990b), and depth (Hinton & Becker, 1990; Marshall, 1989b; Sun, Chen, & Lee, 1987). Unlike supervised learning techniques, such as backpropagation (Rumelhart, Hinton, & Williams, 1986; Werbos, 1974), unsupervised methods have the advantage of not requiring that an external "teacher" be available to program the system; in the lower levels of animal visual systems no external teacher is available (Reeke, Finkel, & Edelman, 1990).

The capability of unsupervised self-organizing neural networks (SONNs) to adapt or learn illustrates how certain visual response properties might be acquired in animal brains, without requiring genetic encoding to carry most of the informational burden of specifying the detailed full cortical wiring plan (Marshall, 1990a; Price & Zumbroich, 1989).

Many self-organizing neural networks (SONNs) have been designed to operate in a winner-take-all fashion. That is, only one neuron or a small, tightly clustered population of neurons is allowed to be active at any moment within a given spatial neighborhood, because of heavy lateral inhibition. Thus any input to such networks is represented as a single, unitary pattern.

Excitatory Learning.

Typically, changes in connection strengths in such networks are governed by a variant of a Hebbian (Hebb, 1949) learning rule, such as the following: *Whenever a neuron is active, its input excitatory connections from active neurons become gradually stronger, while its input excitatory connections from inactive neurons become gradually weaker.* The excitatory learning rule can be expressed mathematically as a differential equation (Grossberg, 1982).

FIGURE 6A depicts the initial architecture of the simplest such network. Two neuron layers are shown, an input layer and a processing layer. In general, such layers can be cascaded, forming a hierarchy of processing stages. A set of initially nonspecific feedforward excitatory connections project from the input layer to the processing layer.

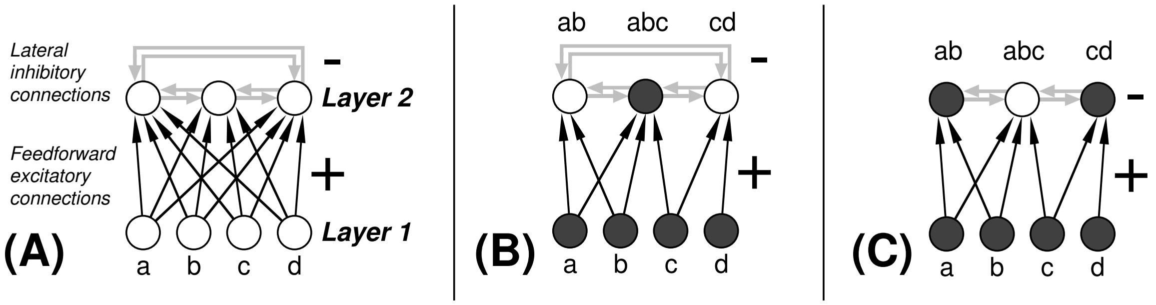

Figure 6: Context-sensitive parsing in an EXIN network. (A) Initially, neurons in the input layer project excitatory connections nonspecifically to neurons in the processing layer. The strength of the excitatory projections may decline as a Gaussian factor of spatial distance, with small random variations. Because each neuron in the processing layer receives approximately the same input connections as its neighbors, the processing-layer neurons are said to have nonspecific receptive fields. In addition, each neuron in the processing layer projects lateral inhibitory connections nonspecifically to all its neighbors (shaded arrows). (B) After a period of exposure to patterns *ab*, *abc*, and *cd*, the excitatory learning rule causes the network to become selective for those patterns: a different neuron becomes wired to respond to each of the familiar patterns. Without inhibitory learning, the network behaves in a winner-take-all fashion: the compound pattern *abcd* (filled lower circles) causes the single "nearest" neuron (*abc*) (filled upper circle) to become active and suppress the activity of the other processing neurons. (C) In an EXIN network, the inhibitory learning rule weakens the strengths of inhibitory connections between neurons that code nonoverlapping patterns, such as between neurons *ab* and *cd*. Then when *abcd* is presented, both neurons *ab* and *cd* become active (filled upper circles), representing the simultaneous presence of the familiar patterns *ab* and *cd*. When pattern *abc* is presented to this EXIN network, neuron *abc* becomes active (not shown in FIGURE) and suppresses the activity of neurons *ab* and *cd*.

Within the processing layer a nonspecific set of lateral inhibitory connections interconnects the neurons. If the lateral inhibitory connections are strong enough, then this network operates as a simple adaptive classifier. When an input pattern is presented (by activating or partially activating some of the input neurons), it most strongly excites the "nearest" processing neuron, which then suppresses the activation of its neighbors in a winner-take-all fashion via lateral inhibition. The feedforward excitatory connection weights to the active processing neuron are then modified slightly to rotate its pattern preferences closer to the input pattern (Amari, 1977; Amari & Takeuchi, 1978; Grossberg, 1976ab; Kohonen, 1982, 1984; Singer, 1983, 1985).

Inhibitory Learning.

Marshall (1990acdf) developed a parsimonious method that allows SONNs to acquire greatly more sophisticated forms of behavior. The technique involves imposing an *inhibitory* learning rule, to govern changes in the strengths of the lateral inhibitory connections, in addition to the excitatory learning rule. Because both excitatory and inhibitory learning rules are used, let us call such SONNs "EXIN" networks.

The EXIN networks use the following inhibitory learning rule, in concert with the excitatory rule described above: *Whenever a neuron is active, its output inhibitory connections to other active neurons become gradually stronger (i.e., more inhibitory), while its output inhibitory connections to inactive neurons become gradually weaker.* The inhibitory rule can also be expressed mathematically as a differential equation (Marshall, 1990acdf).

Thus, if two neurons are frequently coactivated, then the strengths of their reciprocal inhibitory connections tend to increase, thereby decreasing the likelihood that they can become coactivated in the future. On the other hand, if two neurons are rarely coactivated, then the strengths of their reciprocal connections tend to decrease, thereby permitting them to become coactivated on relatively rare occasions. Both the excitatory and the inhibitory learning rules are variants of Hebb's (1949) learning rule; they are simple, physically local, and neurologically plausible.

The earlier networks, with winner-take-all behavior, are capable of representing each input pattern only as a single, lumped item. If an input containing multiple superimposed familiar patterns is presented, such networks are incapable of simultaneously representing all of the patterns separately.

The new EXIN method allows *multiple* patterns to be represented simultaneously, in a distributed but highly regulated fashion. The new unsupervised inhibitory learning rule produces a SONN structure that allows multiple neurons to "win" a network competition, instead of forcing a single winner to "take all." Thus, when multiple independent input patterns are present, all of them can be represented by distinct activations. The excitatory and inhibitory learning rules fulfill different but complementary roles.

While others (Carlson, 1990; Easton & Gordon, 1984; Kohonen, 1984; Földiák, 1989, 1990; Nagano & Kurata, 1981; Nigrin, 1990ab; Rubner & Schulten, 1990) have proposed similar learning rules, Marshall (1990acdef) showed how a network using both an excitatory and an inhibitory learning rule can *multiplex* (Grossberg & Marshall, 1989; Marshall, 1990c), or represent simultaneously, more than one superimposed familiar input pattern.

The end-result of EXIN learning is a network that behaves like the *masking field* networks of Cohen and Grossberg (1986, 1987). However, because the EXIN networks do not require codes for all possible input patterns to be prewired, they are more flexible and efficient than masking fields.

Operation of EXIN Networks.

The *crucial* importance of context sensitivity in visual processing is illustrated by the following example (Nigrin, 1990a). When input pattern *abc* is presented to the network described in the previous paragraph, neuron *abc* becomes active, thereby coding the pattern.

Now suppose instead that occasionally the familiar stimuli *ab* and *cd* are presented simultaneously, so that the input pattern to the network is *abcd*. How is this pattern coded? In a winner-take-all network, the best response is to activate neuron *abc*, because it represents the closest match to the input pattern (FIGURE 6B). However, such a response ignores the presence of *d*, as if *d* were merely noise.

But in an EXIN network, inhibitory strengths between neurons *ab* and *cd* are weak (because *ab* and *cd* do not overlap spatially and hence are not always excited simultaneously), while inhibitory strengths between *ab* and *abc* and between *cd* and *abc* are strong (because of their overlap and consequent frequent simultaneous excitation) (FIGURE 6C). Thus, when *abcd* is presented, neuron *abc* receives inhibition from *both* *ab* and *cd*, while *ab* and *cd* each receive inhibition *only* from *abc*. Because *abc* receives more inhibition, its activation is suppressed, and *both* *ab* and *cd* become active. The simultaneous activation of both *ab* and *cd* is made possible because of the weakened reciprocal inhibition between them.

The EXIN network thus chooses a more complete representation of the input, in which familiar patterns *ab* and *cd* were superimposed, than is possible in the winner-take-all network. The contextual presence of *d* completely alters the multiplexed *parsing* of the input.

One reason that previous research had been unable to successfully extend winner-take-all models to higher levels of visual processing is their relatively poor handling of multiplexing and context.

Properties and Uses of EXIN Networks.

The EXIN rules have the nice dual effects of identifying patterns by frequency of occurrence, or familiarity, and then building a mechanism that chooses a near-optimal representation of multiple superimposed patterns, by simultaneous distributed activation of multiple processing neurons. The behavior of the EXIN networks is superior in this regard to that of ordinary "k-winner" networks, which merely activate the k most-excited processing neurons, regardless of the coactivation or familiarity of the patterns coded by those neurons. Instead, the EXIN networks implement a form of credit assignment (Barto, Sutton, & Anderson, 1983); they can even suppress the activation of some of the most-excited processing neurons when necessary in order to achieve a maximally context-sensitive, global representation of the input. Marshall (1990c) analyzes the details of this process and includes a description of why it works with analog as well as binary patterns and with patterns of different sizes.

The distributed representations allowed in EXIN networks also make them more efficient than k-winner networks, in the following sense: conjunctions (like *abcd*) of familiar patterns (ab, cd) do not require additional neurons to be accurately represented. In order to represent *abcd* as distinct from *abc*, a k-winner network would need an extra processing neuron, devoted to coding *abcd*. In fact, it would need an extra neuron for *every* possible conjunction of input patterns, whereas the EXIN networks can multiplex the coding of conjunctions by simultaneous activation of several neurons. The multiplexed distributed coding in EXIN networks therefore avoids or at least minimizes the "grandmother cell" dilemma of combinatorial explosion.

In vision, EXIN networks perform a context-sensitive parsing, independently representing multiple objects or features at once. The EXIN networks have already been used to model (Marshall, 1990abef) several aspects of visual perception, including the development of disambiguation mechanisms for pattern-motion (Movshon, Adelson, Gizzi, & Newsome, 1985) in the aperture problem (Adelson & Movshon, 1982; Hildreth, 1983; Marr, 1982; Marr & Ullman, 1981), the development of hypercomplex-type length-sensitivity (Hubel & Wiesel, 1977; Kato, Bishop, & Orban, 1978; Orban, Kato, & Bishop, 1979), and the development of orientation sensitivity (Hubel & Wiesel, 1962) in neurons capable of representing edge intersections (Walters, 1987). Preliminary results (Marshall, 1990e) have also been successful in using the EXIN learning rules to model the development of neural representation of binocular disparity surfaces, including stereo transparency (Prazdny, 1985; Weinshall, 1989; Lehky & Sejnowski, 1990).

Development of Lateral Excitatory Connections.

Modeled LEHICs with signal transmission latencies were added to the processing layer of an EXIN network (Marshall, 1989ab, 1990a). Changes in the weights of the connections were governed by the excitatory learning rule. The transmission delays allowed the network to learn to represent input *sequences*; they have been used in learning visual motion sequences (Marshall, 1989ab, 1990a).

An EXIN network using such LEHICs can learn to disambiguate uncertainty in the perception of motion. The network is exposed to visual motion sequences and internalizes a model of the visual inertia characteristics of the motion. That is, it learns that in ordinary visual experience there is a high correlation between detecting motion in a certain direction at a certain location, and detecting motion slightly later, in the same direction but at a location slightly farther along a trajectory in that direction.

The network's knowledge of such natural visual inertia correlations takes the form of a "chain" structure that develops across its LEHICs (FIGURE 7). Each link of the chain corresponds to the correlation between activation of a motion detector at one retinotopic location and activation of another one slightly later, at a successive spatial location.

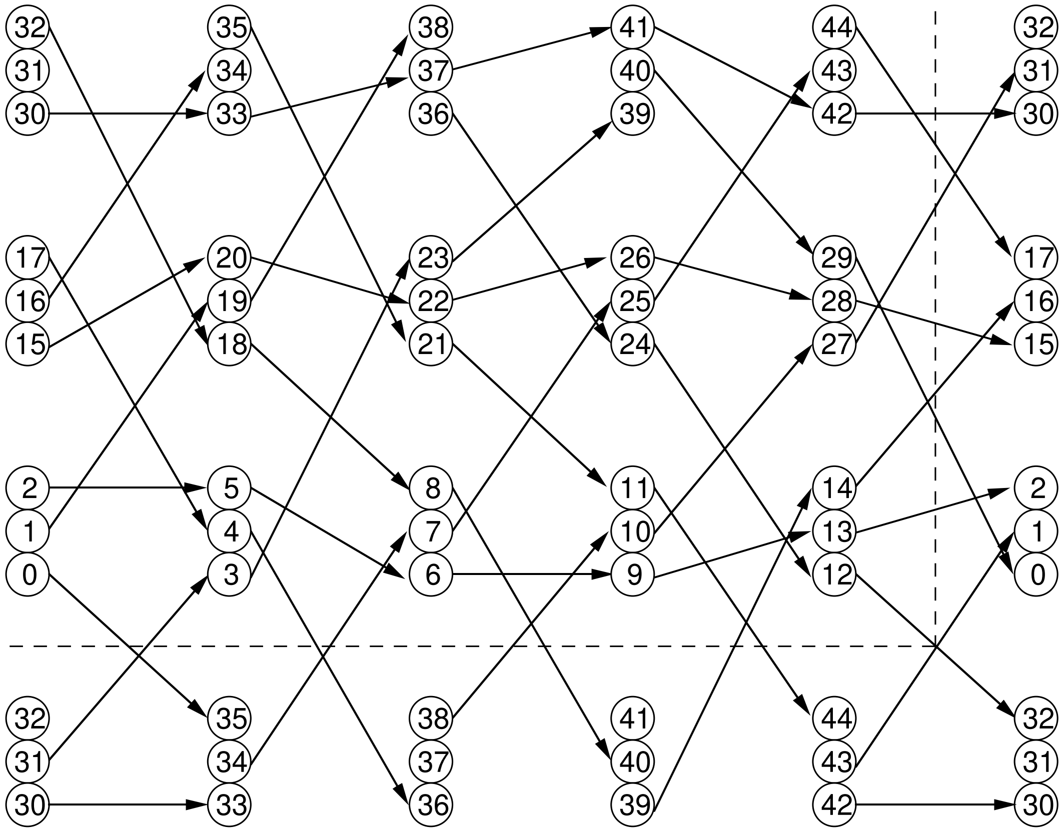

Figure 7: Lateral excitatory connection chains. Time=8000. The chain lattice has formed perfectly. Each neuron in 3×5 matrix of clusters is identified by number. Some neurons are shown more than once to facilitate display of "wraparound" connections, which were used in the simulation to circumvent edge effects. Initially, every neuron in this processing layer was connected weakly to all its neighbors in this layer. After 8000 time steps, all strong remaining LEHICs are part of a lateral chain. Also, each cell in every cluster is part of a *different* chain. The same overall chain structure remains permanently through time 15,000, even though learning is allowed to continue.

The network operates by using the chain connections to track predictively the successive locations of moving stimuli. Local measurement uncertainty in perceived direction of motion (the aperture problem) is resolved through this tracking mechanism (FIGURE 8).

The horizontal intrinsic connections that develop exhibit several relevant characteristics:

- they are highly specific, serving as pathways by which a representation of a particular visual pattern can influence successive representations of the pattern;
- they interconnect neurons that have like stimulus preferences;
- they traverse long ranges of retinotopic distance;
- they serve to predictively prime or disambiguate locally uncertain image measurements, based on the recent history of visual features;

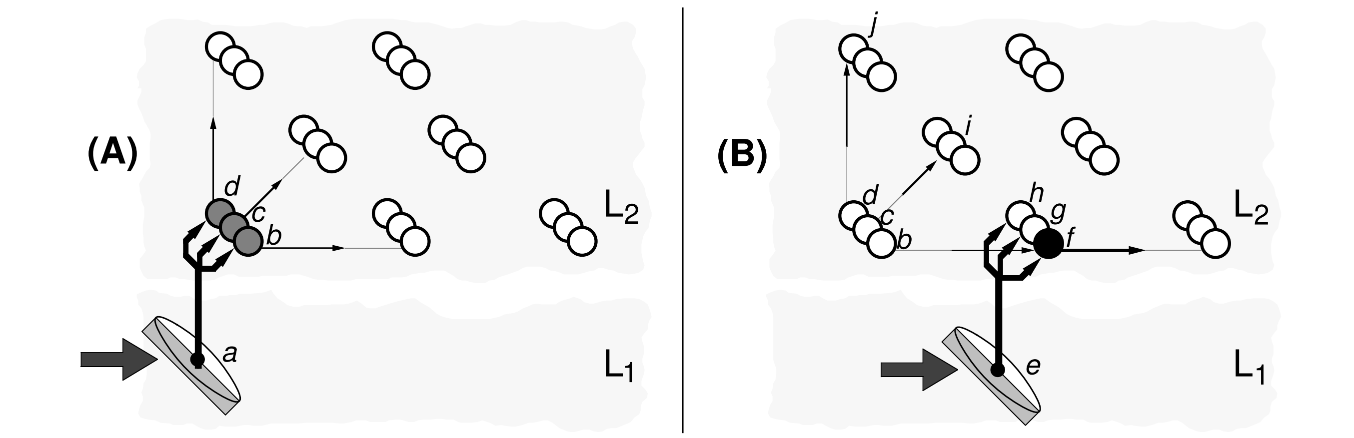

Figure 8: Propagation of Motion Signals along LEHIC Chains. (A) Phase 1: Uncertainty. Horizontally moving edge activates neuron a (a local-motion detector) in layer L_1, which excites neurons b, c, d in layer L_2. Moderately active L_2 neurons (shaded) emit lateral excitatory signals, which do not arrive at their destinations yet because of transmission latencies. The simultaneous activation of the three neurons b, c, d represents the network's uncertainty about the actual direction of motion of the edge. (B) Phase 2: Resolution. A short while later, the edge has moved farther to the right and activates L_1 neuron e (another local-motion detector), which excites L_2 cluster f, g, h. Lateral signals now reach their destinations: neurons f, i, j. One neuron receives both bottom-up plus lateral excitation, becomes strongly active (solid), and suppresses its neighbors' activations via lateral inhibition within its cluster. Neurons that receive only lateral excitation (i, j) are too weakly activated to propagate their lateral signals. Only one neuron per cluster propagates its lateral signals in this phase. The choice of a single active neuron represents the network's decision about the actual direction of motion of the edge.

- they are adaptive – indeed, they acquire all their specificity through observing and internalizing correlations in the visual world.

Based on this work, we may also posit an additional hypothesis:

Development/Adaptation/Learning Hypothesis. Some or all of the visual grouping, steering, and segmentation capabilities described above (in the section on Nonlocal Phenomena in Visual Perception) develop postnatally in human vision after a sufficient period of exposure to the visual world. Certain activity-dependent long-term adaptation processes may directly synthesize the necessary specific predictive pathways, as described below. Much of the anatomically and physiologically observed structure of horizontal intrinsic connections in several regions of visual cortex may be attributable to such activity-dependent development. The pathways identified through neural network simulation of such adaptive processes may be homologous to those found in parts of visual cortex.

Applications of EXIN Networks in Modeling Higher-Level Visual Processing.

In an EXIN network, multiple visual features, objects, or transformations can undergo simultaneous processing, even if they overlap (as in the case of transparency) or are discontinuous (as in the case of occlusion boundaries). The EXIN rules provide both a means by which visual groupings can be learned (EX) and a means by which the learned groupings can be either engaged with or dissociated from one another (IN).

Using EXIN learning rules plus LEHICs can lead to the following capabilities:

- **Invisibility Representation and Imputation**. The LEHICs can provide partial excitation for neurons that lack stimulation from actual incoming, visible image data. This kind of excitation arises when an image feature would be predicted to appear at a certain location but fails to actually appear there because of occlusion or other factors. In theory, the network rules could be designed to let such under-excited neurons propagate such signals, under appropriate conditions and with appropriate restrictions.

- **Grouping**. The selectivity properties of the networks can let different neurons develop receptive fields that are identical except for their lateral input connections. Then a network would have different ways to represent a given feature (i.e., by activating different neurons) (Marshall, 1990a). For each such feature, only one of these representations would receive lateral excitatory signals from previously activated neurons. The lateral signals would select that neuron as the one currently representing the feature; the others' activations would be suppressed. The selection of a single active neuron would represent the *grouping* to which the feature belongs, based on contextual data. The feature would then propagate different output lateral signals, depending on which grouping it belongs to.
- **Coherence**. The inhibitory learning rule lets neurons coding inconsistent attributes (e.g., left motion and right motion) become disconnected from one another, because they are rarely coactivated. Later, if both attributes are present simultaneously (as in motion transparency), then features with one attribute are linked together into one group, and features with the other are linked into another group. Neurons coding one attribute simply ignore features with other attributes. This implements a form of sensitivity to "common fate."
- **Steering**. The rules of neural operation can be designed to give priority to actual incoming visual data, over predicted data coming in via the LEHICs. Thus, visible data can control, or steer, the motion of the representation of an entire group, including invisible features.
- **Trajectory-Specific Propagation**. The LEHICs send signals to positions in the direction and velocity corresponding to each object's predicted trajectory. These signals can bias neuron activations in response to the subsequent incoming image data, to implement a tracking mechanism. The signals are transmitted along different connections for each visual object or feature. Hence, the representation is object-centered; signals pertaining to different objects do not interfere with one another. Besides learning to match the predicted position, direction, and velocity of each object, the LEHICs can in theory also learn to match the predicted configuration and appearance of the same object. A *pattern* of activations across several neurons (rather than a single neuron's activation), corresponding to one object, would send these biasing signals to another pattern of activations.
- **Scission**. The inhibitory learning rule allows representations of different patterns to be parsed and represented simultaneously, even if they overlap spatially.

It is likely that one additional tool will be needed to model some of these perceptual phenomena: top-down excitatory connections, from a network layer to an earlier layer. Like the bottom-up excitatory connections, the top-down connections will obey a variant of a Hebbian excitatory learning rule. This tool may enable specific active groupings to influence the behavior or interpretation of neuron activations at a prior layer.

The EXIN techniques are therefore suitable to be applied to some of the most difficult problems in vision theory, involving nonlocal interactions in visual depth, motion, and form perception. In general, the visual applications of EXIN networks with LEHICs include the following:

Context-sensitive image segmentation. EXIN networks provide a direct means to incorporate feedback of figural information, from representations of objects, into the segmentation process. Thus, segmentation decisions need not be based solely on local measures of discontinuity; instead they can also use information about entire objects, thereby improving segmentation performance.

Transparency (Stereo and Motion). EXIN networks provide a means to form separate groupings (representations of distinct surfaces) within the same 2–D image region. Unlike prior methods, multiple superimposed representations of a limited number can be maintained without significant mutual interference, yet constraints such as uniqueness are preserved.

Efficient Visual Coding. Visual features or objects are coded in EXIN networks according to the learned likelihood of their occurrence in the visual environment. Unlike prior methods, EXIN networks can self-organize yet code inputs containing multiple patterns or ambiguous patterns with significantly less code erosion.

Visual Feature/Object Recognition and Form Representation. Like prior methods, EXIN networks provide a means for low-level feature-detector neurons, analogous to ones found in the visual cortex of many types of animal, to self-organize. But the inhibitory learning capabilities of EXIN networks, combined with learned feedback connections, allow neurons capable of extracting and representing higher-order visual features, including 2–D shape, 3–D shape, depth, and motion-in-depth, to emerge.

Completion of Noisy or Occluded Visual Features. EXIN networks can use feedback from representations of objects to amodally (Day & Kasperczyk, 1983; Kaplan, 1969; Kellman & Cohen, 1984; Michotte, Thinges, & Crabbe, 1964; Parks, 1965) fill-in visual features that are missing or invisible, and to resolve representation ambiguities. This capability is useful in representing partially occluded objects and in maintaining the stability of representations of intermittently occluded objects.

Conclusions: Bag of Tricks or Unified Theory?

A key objective of research in SONNs should be to discover ways in which visual systems can automatically learn to perform a variety of nonlocal operations involved in visual perception of higher-level attributes such as depth, motion, and form. Results along these lines would contribute to an *integrated* theory of self-organizing neural sensorimotor mechanisms, in which multiple processing modules can share a common form of representation, can automatically acquire the necessary structure to perform a variety of tasks, and can communicate effectively with one another.

The challenge of such research is to seek and identify common developmental principles underlying many aspects of vision; the research thus constitutes an attempt to make vision theory more parsimonious. The most significant advantage that a SONN approach to vision may hold over prior approaches is the potential for *unification* of many aspects of visual function under a single common theory. More specifically, the repeated application of a *single* small set of learning rules (perhaps up to parameter differences) at different network loci may generate *many* different visual processing mechanisms. It may become possible to understand the elements of visual systems' diverse "bag of tricks" as different manifestations of a conceptually simple set of self-organization principles. We would then better understand how visual processing areas can represent, communicate, and process information in a self-consistent, rather than piecemeal, manner.

Acknowledgements.

This research was supported by a UNC-CH Junior Faculty Development Award. I thank Pamela K. Johnson for drawing the figures, Bryan Morse, Kevin Martin, and Michael Bradshaw for implementing the motion displays on the computer, and Steven Ornat for helping videotape the displays.

References.

Adelson, E.H. & Movshon, J.A. (1982). "Phenomenal Coherence of Moving Visual Patterns." *Nature, 300*, 523–525.
Amari, S. (1977). "Neural Theory of Association and Concept Formation." *Biological Cybernetics, 26*, 175–185.
Amari, S. & Takeuchi, A. (1978). "Mathematical Theory on Formation of Category Detecting Nerve Cells." *Biological Cybernetics, 29*, 127–136.
Anstis, S.M. & Ramachandran, V.S. (1987). "Visual Inertia in Apparent Motion." *Vision Research, 27*, 5, 755–764.
Barlow, H.B. (1979). "Reconstructing the Visual Image in Space and Time." *Nature, 279*, 189–190.
Barlow, H.B. (1981). "Critical Limiting Factors in the Design of the Eye and Visual Cortex." *Proceedings of the Royal Society of London, Ser. B, 212*, 1–34.
Barto, A.G., Sutton, R.S., & Anderson, C.W. (1983). "Neuronlike Adaptive Elements that Can Solve Difficult Learning Control Problems." *IEEE Transactions on Systems, Man, and Cybernetics, 13*, 834–846.
Bienenstock, E.L., Cooper, L.N., & Munro, P.W. (1982). "Theory for the Development of Neuron Selectivity: Orientation Specificity and Binocular Interaction in Visual Cortex." *Journal of Neuroscience, 2*, 1, 32–48.
Blasdel, G.G., Lund, J.S., & Fitzpatrick, D. (1985). "Intrinsic Connections of Macaque Striate Cortex: Axonal Projections of Cells Outside Lamina 4C." *Journal of Neuroscience, 5*, 12, 3350–3369.
Burr, D. & Ross, J. (1986). "Visual Processing of Motion." *Trends in Neuroscience, 9*, 7, 304–307.
Callaway, E.M. & Katz, L.C. (1990). "Emergence and Refinement of Clustered Horizontal Connections in Cat Striate Cortex." *Journal of Neurophysiology, 10*, 1134–1153.
Carlson, A. (1990). "Anti-Hebbian Learning in a Non-Linear Neural Network." *Biological Cybernetics, 64*, 171–176.
Cohen, M.A. & Grossberg, S. (1986). "Neural Dynamics of Speech and Language Coding: Developmental Programs, Perceptual Grouping, and Competition for Short Term Memory." *Human Neurobiology, 5*, 1–22.
Cohen, M.A. & Grossberg, S. (1987). "Masking Fields: A Massively Parallel Neural Architecture For Learning, Recognizing, and Predicting Multiple Groupings of Patterned Data." *Applied Optics, 26*, 1866–1891.

Coolen, A.C.C. & Kuijk, F.W. (1989). "A Learning Mechanism for Invariant Pattern Recognition in Neural Networks." *Neural Networks, 2*, 495–506.

Day, R.H. & Kasperczyk, R.T. (1983). "Amodal Completion as a Basis for Illusory Contours." *Perception & Psychophysics, 33*, 4, 355-364.

Easton, P. & Gordon, P.E. (1984). "Stabilization of Hebbian Neural Nets by Inhibitory Learning." *Biological Cybernetics, 51*, 1–9.

Földiák, P. (1989). "Adaptive Network for Optimal Linear Feature Extraction." *Proceedings of the International Joint Conference on Neural Networks,* Washington, DC, June 1989, *I.*, 401–405.

Földiák, P. (1990). "Forming Sparse Representations by Local Anti-Hebbian Learning." *Biological Cybernetics, 64*, 165–170.

Földiák, P. (1991). "Learning Invariance from Transformation Sequences." *Neural Computation, 3*, 194–200.

Gabbott, P.L.A., Martin, K.A.C., & Whitteridge, D. (1987). "Connections Between Pyramidal Neurons in Layer 5 of Cat Visual Cortex (Area 17)." *Journal of Comparative Neurology, 259*, 364–381.

Gilbert, C.D. & Wiesel, T.N. (1989). "Columnar Specificity of Intrinsic Horizontal and Corticocortical Connections in Cat Visual Cortex." *Journal of Neuroscience, 9*, 2432–2442.

Grossberg, S. (1976a). "Adaptive Pattern Classification and Universal Recoding: I. Parallel Development and Coding of Neural Feature Detectors." *Biological Cybernetics, 23*, 121–134.

Grossberg, S. (1976b). "On the Development of Feature Detectors in the Visual Cortex with Applications to Learning and Reaction-Diffusion Systems." *Biological Cybernetics, 21*, 145–159.

Grossberg, S. (1982). *Studies of Mind and Brain: Neural Principles of Learning, Perception, Development, Cognition, and Motor Control.* Boston: Reidel Press.

Grossberg, S. & Marshall, J.A. (1989). "Stereo Boundary Fusion by Cortical Complex Cells: A System of Maps, Filters, and Feedback Networks for Multiplexing Distributed Data." *Neural Networks, 2*, 29–51.

Grossberg, S. & Mingolla, E. (1985a). "Neural Dynamics of Form Perception: Boundary Completion, Illusory Figures, and Neon Color Spreading." *Psychological Review, 92*, 173–211.

Grossberg, S. & Mingolla, E. (1985b). "Neural Dynamics of Perceptual Grouping: Textures, Boundaries, and Emergent Segmentations." *Perception & Psychophysics, 38*, 141-171.

Hebb, D.O. (1949). *The Organization of Behavior.* New York: Wiley.

Hildreth, E.C. (1983). "Computing the Velocity Field Along Contours." *Proceedings of the ACM SIGGRAPH/SIGART Interdisciplinary Workshop on Motion*, Toronto: 26–32.

Hinton, G.E. & Becker, S. (1990). "An Unsupervised Learning Procedure that Discovers Surfaces in Random-Dot Stereograms." *Proceedings of the International Joint Conference on Neural Networks*, Washington DC, January 1990, *I.*, 218–222.

Hubel, D.H. & Wiesel, T.N. (1962). "Receptive Fields, Binocular Interactions, and Functional Architecture in Cat's Visual Cortex." *Journal of Physiology, 160*, 106–154.

Hubel, D.H. & Wiesel, T.N. (1977). "Functional Architecture of Macaque Monkey Visual Cortex." *Proceedings of the Royal Society of London, Ser. B, 198*, 1–59.

Kanizsa, G. (1979). *Organization in Vision: Essays on Gestalt Perception.* New York: Praeger.

Kaplan, G.A. (1969). "Kinetic Disruption of Optical Texture: The Perception of Depth at an Edge." *Perception & Psychophysics, 6*, 4, 193–198.

Kato, H., Bishop, P.O., & Orban, G.A. (1978). "Hypercomplex and Simple/Complex Cell Classifications in Cat Striate Cortex." *Journal of Neurophysiology, 41*, 1071–1095.

Kellman, P.J. & Cohen, M.H. (1984). "Kinetic Subjective Contours." *Perception & Psychophysics, 35*, 3, 237–244.

Kohonen, T. (1982). "Self-Organized Formation of Topologically Correct Feature Maps." *Biological Cybernetics, 43*, 59–69.

Kohonen, T. (1984). *Self-Organization and Associative Memory.* New York: Springer-Verlag.

Lehky, S.R. & Sejnowski, T.J. (1990). "Neural Model of Stereoacuity and Depth Interpolation Based on a Distributed Representation of Stereo Disparity." *Journal of Neuroscience, 10*, 7, 2281–2299.

Linsker, R. (1986a). "From Basic Network Principles to Neural Architecture: Emergence of Spatial-Opponent Cells." *Proceedings of the National Academy of Sciences of the U.S.A., 83*, 7508–7512.

Linsker, R. (1986b). "From Basic Network Principles to Neural Architecture: Emergence of Orientation-Selective Cells." *Proceedings of the National Academy of Sciences of the U.S.A., 83*, 8390–8394.

Luhmann, H.J., Martínez Millán, L., & Singer, W. (1986). "Development of Horizontal Intrinsic Connections in Cat Striate Cortex." *Experimental Brain Research, 63*, 443–448.

Lund, J.S. (1987). "Local Circuit Neurons of Macaque Monkey Striate Cortex: I. Neurons of Laminae 4C and 5A." *Journal of Comparative Neurology, 257*, 60–92.

Marr, D. (1982). *Vision.* San Francisco: W.H. Freeman and Company.

Marr, D. & Poggio, T. (1976). "Cooperative Computation of Stereo Disparity." *Science, 194*, 283–287.

Marr, D. & Ullman, S. (1981). "Directional Selectivity and Its Use in Early Visual Processing." *Proceedings of the Royal Society of London, Ser. B, 211*, 151–180.

Marshall, J.A. (1989a). *Neural Networks for Computational Vision: Motion Segmentation and Stereo Fusion. Ph.D. Dissertation, Boston University.* Ann Arbor, Michigan: University Microfilms Inc.

Marshall, J.A. (1989b). "Self-Organizing Neural Network Architectures for Computing Visual Depth from Motion Parallax." *Proceedings of the International Joint Conference on Neural Networks*, Washington DC, June 1989, *II.*, 227–234.

Marshall, J.A. (1990a). "Self-Organizing Neural Networks for Perception of Visual Motion." *Neural Networks, 3*, 45–74.

Marshall, J.A. (1990b). "Development of Length-Selectivity in Hypercomplex-Type Cells." *Investigative Ophthalmology and Visual Science, 31*, 4, 397.

Marshall, J.A. (1990c). "A Self-Organizing Scale-Sensitive Neural Network." *Proceedings of the International Joint Conference on Neural Networks*, San Diego, June 1990, *III.*, 649–654.

Marshall, J.A. (1990d). "Representation of Uncertainty in Self-Organizing Neural Networks." *Proceedings of the International Neural Network Conference*, Paris, France, July 1990, 809–812.

Marshall, J.A. (1990e). "Self-Organizing Neural Network for Computing Stereo Disparity and Transparency." *Optical Society of America Annual Meeting Technical Digest*, Boston, November 1990, 268.

Marshall, J.A. (1990f). "Adaptive Neural Networks for Multiplexing Oriented Edges." In D.P. Casasent (Ed.), *Intelligent Robots and Computer Vision IX: Neural, Biological, and 3-D Methods*, *Proceedings of the SPIE 1382*, Boston, November 1990, pp.282–291.

Michalski, A., Gerstein, G.L., Czarkowska, J., & Tarnecki, R. (1983). "Interactions Between Cat Striate Cortex Neurons." *Experimental Brain Research, 51*, 97–107.

Michotte, A., Thinges, G., & Crabbe, G. (1964). "Les Complements Amodaux des Structures Perceptives (Studia Psychologia)." Louvain: Publications Universitaires de Louvain.

Mitchison, G. & Crick, F. (1982). "Long Axons Within the Striate Cortex: Their Distribution, Orientation, and Patterns of Connection." *Proceedings of the National Academy of Sciences of the U.S.A., 79*, 3661–3665.

Movshon, J.A., Adelson, E.H., Gizzi, M.S., & Newsome, W.T. (1985). "The Analysis of Moving Visual Patterns." In C. Chagas, R. Gattass, & C. Gross (Eds.), *Pattern Recognition Mechanisms.* Vatican City: Pontifical Academy of Sciences, 117–151.

Nagano, T. & Kurata, K. (1981). "A Self-Organizing Neural Network Model for the Development of Complex Cells." *Biological Cybernetics, 40*, 195–200.

Nakayama, K., Shimojo, S., & Silverman, G.H. (1989). "Stereoscopic Depth: Its Relation to Image Segmentation, Grouping, and the Recognition of Occluded Objects." *Perception, 18*, 55–68.

Nelson, J.I. & Frost, B.J. (1985). "Intracortical Facilitation Among Co-Oriented, Co-Axially Aligned Simple Cells in Cat Striate Cortex." *Experimental Brain Research, 61*, 54–61.

Nigrin, A.L. (1990a). "SONNET: A Self-Organizing Neural Network that Classifies Multiple Patterns Simultaneously." *Proceedings of the International Joint Conference on Neural Networks*, San Diego, June 1990, *II.*, 313–318.

Nigrin, A.L. (1990b). *The Stable Learning of Temporal Patterns with an Adaptive Resonance Circuit. Ph.D. Dissertation, Duke University.*

Orban, G.A., Kato, H., & Bishop, P.O. (1979). "End-Zone Region in Receptive Fields of Hypercomplex and Other Striate Neurons in the Cat." *Journal of Neurophysiology, 42*, 3, 818–832.

Parks, T. (1965). "Post-Retinal Visual Storage." *American Journal of Psychology, 78*, 145–147.

Parks, T. (1986). "Illusory Figures, Illusory Objects, and Real Objects." *Psychological Review, 93*, 207–215.

Pasternak, T. & Leinen, L.J. (1986). "Pattern and Motion Vision in Cats with Selective Loss of Cortical Directional Selectivity." *Journal of Neuroscience, 6*, 4, 938–945.

Piaget, J. (1952). *The Origins of Intelligence in Children.* New York: International University Press.

Prazdny, K. (1985). "Detection of Binocular Disparities." *Biological Cybernetics, 52*, 93–99.

Prazdny, K. (1986). "Illusory Contours from Inducers Defined Solely by Spatiotemporal Correlation." Preprint.

Price, D.J. & Zumbroich, T.J. (1989). "Postnatal Development of Corticocortical Efferents from Area 17 in the Cat's Visual Cortex." *Journal of Neuroscience, 9*, 2, 600–613.

Ramachandran, V.S., Cobb, S., & Rogers-Ramachandran, D. (1988). "Perception of 3-D Structure from Motion: The Role of Velocity Gradients and Segmentation Boundaries." *Perception & Psychophysics, 44*, 4, 390–393.

Ramachandran, V.S., Inada, V., & Kiama, G. (1986). "Perception of Illusory Occlusion in Apparent Motion." *Vision Research, 26*, 10, 1741–1749.

Reeke, G.N., Finkel, L.H., & Edelman, G.M. (1990). "Selective Recognition Automata." In S.F. Zornetzer, J.L. Davis, & C. Lau (Eds.), *An Introduction to Neural and Electronic Networks*, New York: Academic Press, 203–226.

Rockland, K.S. & Lund, J.S. (1982). "Widespread Periodic Intrinsic Connections in the Tree Shrew Visual Cortex." *Science, 215*, 1532–1534.

Rockland, K.S. & Lund, J.S. (1983). "Intrinsic Laminar Lattice Connections in Primate Visual Cortex." *Journal of Comparative Neurology, 216*, 303–318.

Rockland, K.S., Lund, J.S., & Humphrey, A.L. (1982). "Anatomical Banding of Intrinsic Connections in Striate Cortex of Tree Shrews (*Tupaia glis*)." *Journal of Comparative Neurology, 209*, 41–58.

Rubner, J. & Schulten, K. (1990). "Development of Feature Detectors by Self-Organization." *Biological Cybernetics, 62*, 193–199.

Rumelhart, D., Hinton, G., & Williams, R. (1986). "Learning Representations by Backpropagating Errors." *Nature, 323*, 533-536.

Sereno, M.E. (1986). "Neural Network Model for the Measurement of Visual Motion." *Journal of the Optical Society of America A, 3*, 13, P72.

Sereno, M.E. (1987). "Implementing Stages of Motion Analysis in Neural Networks." *Program of the Ninth Annual Conference of the Cognitive Science Society*, Hillsdale, NJ: Lawrence Erlbaum Associates, 405–416.

Sereno, M.E., Kersten, D.J., & Anderson, J.A. (1988). "A Neural Network Model of an Aspect of Motion Perception." *Science at the John von Neumann National Supercomputer Center: Annual Report FY 1988*, 173–178.

Sereno, M.I. (1989). "Learning the Solution to the Aperture Problem for Pattern Motion with a Hebb Rule." In D. Touretzky (Ed.), *Advances in Neural Information Processing Systems, 2*, San Mateo, CA: Morgan Kaufman.

Sereno, M.I. & Sereno, M.E. (1990). "Learning to See Rotation and Dilation with a Hebb Rule." To appear in D. Touretzky (Ed.), *Advances in Neural Information Processing Systems, 3*, San Mateo, CA: Morgan Kaufman.

Shimojo, S., Silverman, G.H., & Nakayama, K. (1988). "An Occlusion-Related Mechanism of Depth Perception Based on Motiona and Interocular Sequence." *Nature, 333*, 265–268.

Singer, W. (1983). "Neuronal Activity as a Shaping Factor in the Self-Organization of Neuron Assemblies." In E. Basar, H. Flohr, H. Haken, & A. J. Mandell (Eds.), *Synergetics of the Brain.* New York: Springer-Verlag.

Singer, W. (1985). "Activity-Dependent Self-Organization of the Mammalian Visual Cortex." In D. Rose & V.G. Dobson (Eds.), *Models of the Visual Cortex.* New York: John Wiley and Sons Ltd., 123–136.

Sun, G.Z., Chen, H.H., & Lee, Y.C. (1987). "Learning Stereopsis with Neural Networks." Preprint.

Ts'o, D.Y., Gilbert, C.D., & Wiesel, T.N. (1986). "Relationships Between Horizontal Interactions and Functional Architecture in Cat Striate Cortex as Revealed by Cross-Correlation Analysis." *Journal of Neuroscience, 6*, 4, 1160–1170.

Ullman, S. (1979). *The Interpretation of Visual Motion.* Cambridge, MA: M.I.T. Press.

Van Essen, D.C. & Anderson, C.H. (1987). "Reference Frames and Dynamic Remapping Processes in Vision." In E. Schwartz (Ed.), *Computational Neuroscience.* Cambridge, MA: MIT Press.

von der Heydt, R., Peterhans, E., & Baumgartner, G. (1984). "Illusory Contours and Cortical Neuron Responses." *Science, 224*, 1260–1262.

von der Malsburg, C. (1973). "Self-Organization of Orientation Sensitive Cells in the Striate Cortex." *Kybernetik, 14*, 85–100.

Walters, D. (1987). "Rho-Space: A Neural Network for the Detection and Representation of Oriented Edges." *Proceedings of the First International Conference on Neural Networks*, San Diego, June 1987.

Weinshall, D. (1989). "Perception of Multiple Transparent Planes in Stereo Vision." *Nature, 341*, 737–739.

Werbos, P. (1974). "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences." *Ph.D. Dissertation, Harvard University.*