The perception of form and motion

Patrick Cavanagh

Harvard University, Cambridge, USA

Although form and motion are two distinct aspects of visual processing, they do not start as separate entities in the visual system. Early analyses extract discontinuities in various image attributes and these can trace the outline of a form. When displaced, the same image features can give rise to impressions of motion. Recent work has overturned many of the assumptions about the contributions of different stimulus attributes to motion processing, and reorganized the classification of motion systems. The results have revealed unexpected interactions between attention and motion. Paralleling this research is work on the early stages of form learning and on the nature of the stored representations.

Current Opinion in Neurobiology 1993, **3**:177–182

Introduction

Research over the past decade has suggested that cortical vision systems have two goals, one concerned with extracting an object's form and the other with tracking its position and motion. These general tasks have been linked to the dorsal and ventral streams in the physiological literature. Preceding these output streams, however, are several earlier streams of analysis more closely tied to stimulus properties of color, texture, motion, depth, and brightness. Recent studies show that any or all of these properties can contribute to the definition of stimulus contours and their motion. A higher level motion system has also been identified that is involved in tracking visible form with attention. In a related development, interactions between attention and motion have been used to map the displacement of attention that occurs when a stimulus moves. There are also important new directions emerging in research on form recognition. Several studies point to early view-specific two-dimensional matches as a basic stage in form recognition, and long-term training experiments have revealed an early pattern learning that is retinotopic and task specific.

Contribution of different attributes to motion

Stimuli used in motion experiments have increasingly been classified into two categories. First-order stimuli (luminance or color) can be defined at a single point. Second-order stimuli are defined by differences in texture, depth, or motion, and require two points to define the local values of each attribute. The two points are separated in space to specify each texture, separated by eye to define each value of binocular disparity, or separated by space and time to define each local value of motion. This classification places color with luminance as a first-order motion stimulus even though early motion processes have traditionally been considered to be color-blind. Recent work shows that this view was based on the use of inappropriate scales for comparing color and luminance stimuli. Using appropriate contrast scales (based on cone response [1••]), I have [2] reported that the contrast threshold for discriminating the direction of motion can be similar for both color and luminance gratings. The often demonstrated loss of performance for chromatic stimuli, which led to the assumption of weak color input to motion, was actually a consequence of the weakness of the stimuli; experiments kept stimulus contrast equivalent for chromatic and luminance stimuli, typically in terms of phosphor modulation, but had not equated cone responses. When this is done, the response of motion processes to chromatic stimuli is strong. Pappathomas, Gorea and Julesz [3] have also reported strong motion responses for color stimuli in an apparent motion display.

Second-order stimulus attributes have also been shown to produce significant motion responses. Stimuli such as texture-defined [4] or stereo-defined gratings [5••,6] can be seen to move. Stereo-defined and luminance-defined stimuli produce very similar results in tests of apparent motion (the Ternus display [5••]). Many aspects of motion processing differ for first- and second-order stimuli, however. Pantle [7] has reported that the response to moving second-order stimuli degrades rapidly in the periphery. Mather [8] has shown that although second-order motion stimuli can produce motion after-effects, these last only half as long as those for first-order stimuli. Werkhoven, Sperling and Chubb [9•] have used a new and powerful ambiguous motion display to show that second-order motion strength between two elements is not based on the similarity between the elements, but

---

Abbreviations

2D—two dimensional; **3D**—three dimensional.

rather covariance in contrast. They propose a simple preprocessing stage followed by 'standard' motion energy analysis to account for second-order motion.

The motion responses to different stimulus attributes appear to be combined in some common pathway. In particular, opposing patterns defined by different attributes cancel each other's motion. For example, gratings defined by color and luminance drifting in opposite directions can produce a motion null so that only an impression of flicker remains [10•]. Similarly, gratings defined by texture and luminance [11] also produce a motion null when superimposed and set in motion in opposite directions. If there were independent analyses of motion for these different attributes, a transparency effect would be expected — the two directions of motion should be visible simultaneously.

Turano [12•] has reported transfer of direction-specific threshold elevation between first-order patterns defined by luminance and second-order patterns defined by contrast-modulated textures. She has concluded that these two pattern types are processed by common mechanisms. Furthermore, Stoner and Albright [13] have shown that motion signals arising from different attributes, such as luminance and texture, can combine to produce a percept of a coherently moving pattern. Their results support earlier physiological findings that the directional selectivity of many neurons in the cortical MT area is unaffected by changes in the attribute that defines the moving stimulus.

Plaids as second-order stimuli

When a vertical and a horizontal grating are superimposed and each set into motion orthogonal to its orientation, a new percept is generated of a rigid 'plaid' pattern that moves in a diagonal direction. This stimulus has often been used to study the aperture problem [14]: the question of how different local motion signals are combined to recover the global motion of an object. Adelson and Movshon's original idea was that the motions of the two component gratings were sensed independently, and then an 'intersection of constraints' computation recovered the motion of the combined pattern. This model received support in many subsequent papers. Two recent reports by Derrington and colleagues have challenged this two-stage model, however. In the first report they have shown that the velocity threshold for discriminating the direction of the plaid motion can be lower than for either of its two components [15]. In the second, they report that when a plaid jumps through a step that moves both of the components by an unambiguous 135°, the plaid is seen to move in the opposite direction [16•]. Apparently the plaid is seen as a textured grating (a second-order stimulus) with alternating bands of high contrast stripes and uniform gray. This emergent grating has double the spatial frequency of the plaid (half the spatial period, alternate bands have shifted positions of light and dark stripes, but this is not immediately noticeable). A 135°shift of the plaid to the right is therefore a 270°shift of this frequency-doubled texture pattern, and the whole image moves clearly 90°to the left. Work by Yo and Wilson [17] also underlines the importance of this texture-based or second-order representation of the plaid in determining its overall motion. They claim that the contributions of the second-order and first-order motion signals combine in a common response mechanism.

Mapping attention with motion

Stelmach and Herdman [18] have shown that attention can influence the apparent timing of events. If attention is directed to a particular location, a subsequent brief flash at that location appears to occur earlier than an identical simultaneous flash placed elsewhere. This express perception of events in an attended location also influences the perception of a larger shape that is partially in the attended region. Hikosaka, Miyauchi and Shimojo [19••] have shown that when a line is flashed so that one end is in an attended region, the line appears to 'turn on' from that end producing an impression of motion along the line, like a fuse being lit. This motion occurs even though all parts of the line are physically displayed at the same time. They have used this phenomenon of induced line motion to probe the distribution of attention in other tasks, in particular, during apparent motion.

Attention-based motion

Although previous studies had proposed two motion systems based on the distance over which motion percepts could be supported — short for fine textures and long for large targets — a recent article [20••] suggests that a different dichotomy may be more appropriate. Specifically, low-level detectors are available to signal stimulus displacement for a variety of first- or second-order stimulus attributes. All of these operate in a similar fashion with a dense array of detectors covering the visual field and analyzing image features in parallel. A second motion system is available that relies on attentive tracking. When an object moves through the visual field, an observer can track it with attention without engaging any eye movements. The motion of the object can be signaled by low-level detectors, but there may also be a second source of motion signals that relies on the displacements of the focus of attention. In a study that suppressed low-level motion signals from a color grating, observers were able to see the bars of the grating move, but only when they were tracked with attention. This second, attention-based system may also play a significant role in resolving correspondence in apparent motion displays, where one set of items is replaced with a second set.

Tracking with attention

While often attracted automatically to moving stimuli, attention may also be voluntarily assigned to particular stimuli to keep track of them. For example, in a display of several moving targets, attention may follow a given subset while ignoring the rest [21,22]. In a typical study of multiple target tracking, 10 or more identical disks move randomly about the display. To begin tracking, a subset of the disks is identified by flashing. Once the flashing stops, the observers must continue to track the disks that had been flashing until a probe occurs a few seconds later. During the tracking period, all the disks are physically identical. Typically, observers can accurately track four or five targets for short periods of time. How do they do this? Pylyshyn and Storm [21] argue that an index is set up for each target to keep track of its position. Yantis [22] has studied the role of grouping in the ability to track multiple targets, and claims that the observer forms a non-rigid polygon with the tracked targets and uses this form as a unitary placekeeper for all the targets.

This tracking function of attention has also been studied by Kahneman, Treisman and Gibbs [23••]. They introduce the concept of an object file as a temporary placekeeper for objects that appear in the scene. Rather than representing the occurrence of the object through some activation of nodes in a long-term memory, they suggest that a separate memory system with a limited number of object files might be used. When the object appears, its position, class and other attributes would be noted in a newly opened file and these would be updated as the object, say a bird, changes position or shape (flaps its wings). According to these authors, the information in the object file would facilitate the classification of a second appearance of the object. Their experiments support this conjecture in a variety of situations.

Learning form from motion

Motion has a special role among visual sensations because its presence carries a fundamental message. The message is the continuity of an object's identity across time and space. The impression of motion created by a moving object signals that the various instances of the shape seen over a sequence of times and locations all belong to one object. In that sense, motion is a primitive pattern recognition system, which matches patterns over time and signals motion if the match is adequate. The object may be changing shape or color as it moves, and may even, in the hands of clever experimenters, be different objects, but the impression of motion indicates that the spatio-temporal pattern is adequate to be considered a single object, at least from the point of view of the motion system. Why should the motion system be so accepting? Evidently, spatiotemporal proximity is a very robust cue for indicating the displacement of a single object, or at least it was before the advent of movies, tachistoscopes and computer graphics. Although this rough matching may be easily misled by modern animation sequences, it may serve a basic function in pattern learning: the role of teacher. As any object moves through our field of view, we are presented with several different viewpoints for the object, and if it is non-rigid, e.g. like a bird or a leopard, we are also given several shape variants for the object. The fact that all of these variants and viewpoints are linked by motion should alert the pattern system to associate or group them together as views of one object. Learning variants of objects is not the only task, or even the first task faced by the visual system. According to Ahumada [24••], a highly related primary task is to compensate for the irregular positioning of receptors on the retina, as well as for the randomness in subsequent wiring. The consequence of this irregularity is that an object's internal representation will change and distort as its image moves over the retina. Ahumada has shown how neural networks can be trained to regularize the randomness in the receptor arrays by associating a fixed output pattern with the input response arising from a given object as it moves through the visual field. Ahumada assumes that the object does not change shape as it moves. It is not a large step to extend this procedure so that the neural network also learns the variants of objects.

Perceptual learning

Over the past two decades, a number of studies have shown that extended practice can dramatically affect performance in visual tasks [25–28]. Recent studies have revealed that there is a great deal of specificity in this practice effect. Karni and Sagi [29••] have trained subjects on a texture segmentation task, where the target patch appears in only one quadrant during training. They found that the improvement in performance was transferred to displays only in that quadrant, and only when the same background texture elements were used. In addition, transfer was found only when tested on the practiced eye (there was no interocular transfer), and the target texture elements could be changed without losing the practice effect. The specificity to retinotopic location and to background structure suggests that practice affects background grouping processes at very early levels in the visual system, such as area V1 where there are still monocularly driven units. Shiu and Pashler [30] also report that the effects of training in an orientation discrimination task do not transfer to untrained visual quadrants. Moreover, there was no effect of practice when the observers attended to the brightness of the lines rather than their orientation. In other words, the practice effect depended on attentive processing of the stimulus orientation. Poggio, Fahle, and Edelman [31] report a practice effect following brief training on a hyperacuity task. The improvement with practice in this case was specific to the orientation of the vernier lines. Nakayama and Shimojo [32] have suggested that perceptual learning underlies the three-dimensional (3D) interpretation of surfaces. They propose a generic view principle that predicts which of several possible interpretations of an image is most likely to be seen. Generic or common views of cubes, for example, have three visible faces, whereas an accidental view of a cube might reveal two or even only one face. Because of the much higher frequency of three-faced views of cubes from the various vantage points experienced over the lifespan, the authors argue that this relation between image structure (three faces) and 3D surfaces (cube) can be, and is learned, by the visual system. Consequently, one-faced views of cubes lead to percepts of squares, not cubes, and three visible faces are required for the percept of an unambiguous cube.

Is form recognition based on multiple, stored two-dimensional views?

Many object recognition theories propose that an internal 3D model is constructed and stored for learned objects. This would provide view-independent recognition when the object is encountered again. More recent studies propose, to the contrary, that multiple two-dimensional (2D) views are stored for each object [33], and that recognition is achieved by matching to the most similar 2D view, or even by interpolation between the appropriate neighboring views [34,35•]. In Bülthoff and Edelman's [35•] experiments, observers were trained on two different views of an amorphous 3D object. In a test phase, observers had to discriminate the trained object from new ones, when the test objects were presented with views falling within, or outside the range of orientations between the two trained objects. The results show that performance drops dramatically outside the range between the two trained views, but remains respectable for test orientations falling between them. The authors argue that these results are incompatible with internal 3D representations, which should have provided reasonable recognition over a wide range of test orientations. The results support a recognition process that can recognize new views by interpolation, only if they fall between the orientations of stored 2D views.

A study of recognition for shadowed familiar figures (e.g. Mooney faces, [33]) shows that recognition is possible even though no 3D representation can be constructed from image data. For theories proposing 3D internal models as the basis for recognition, the first step is to construct a 3D interpretation of the image and then match it against stored 3D models. For the shadowed test objects, this initial step is impossible because the structure of its shadows cannot be interpreted unless it is already known that the image is, say, a face. These images are nevertheless readily recognized. The only way to gain the initial knowledge that the image is a face is by 2D matching of characteristic face contours embedded in the image. This crude first match could then guide further interpretation of the image data.

Conclusions

The analyses of motion mechanisms have recently focused on second-order stimuli and on interactions between motion and attention. Color itself has been shown to be a strong contributor to motion processes despite early doubts based on inappropriate scales for comparing color and luminance stimuli. It is now known that appropriate scales are based on cone contrast. Second-order stimuli (moving texture-, motion- or stereo-defined contours) produce reasonable motion responses, but appear less robust than those based on moving luminance contours. Experimental results suggest that all of these motion responses combine in a final common pathway. A second motion system based on attentive tracking has been identified and appears to account for many of the phenomena previously classified as 'long-range,' but may be limited to processing four or five targets at any one time. Perceptual learning is currently attracting significant interest and may involve modifications in the interconnections of early vision. Learning can be demonstrated following brief or extended practice on specific tasks. The learning effects are often restricted to the trained retinal location, eye, and stimulus elements and may reveal the characteristics of early visual processing. Recent studies argue that object recognition may be based on multiple 2D views of each object rather than stored 3D models.

Acknowledgements

Preparation of this review was supported by NEI EY09258 and by AFOSR 0169.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. •• CHAPARRO A, STROMEYER III CF, HUANG EP, KRONAUER RE, ESKEW RT JR: **Colour is What the Eye Sees Best.** *Nature* 1993, 361:348–350.

Detection thresholds reveal that the best detected color stimulus is seen three to eight times better than the best detected luminance stimulus. These results are consistent with the prevalence and high color contrast gain of color-opposed retinal P cells, and may compensate for the low chromatic contrasts typically found in natural scenes. The contrasts for color and luminance stimuli are both expressed in the common metric of the modulation of cone output.

2. CAVANAGH P: **The Contribution of Colour to Motion.** In *From Pigments to Perception*. Edited by Valberg A, Lee BB. New York: Plenum Press; 1991:151–164.

3. PAPATHOMAS TV, GOREA A, JULESZ B: **Two Carriers for Motion Perception: Color and Luminance.** *Vision Res* 1991, 31:1883–1891.

4. CHUBB C, SPERLING G: **Texture Quilts: Basic Tools for Studying Motion-from-Texture.** *J Math Psychol* 1991, 35:411–442.

5. •• PATTERSON R, HART P, NOWAK D: **The Cyclopean Ternus Display and the Perception of Element Versus Group Movement.** *Vision Res* 1991, 31:2085–2093.

A bistable stroboscopic motion (Ternus) display was examined as a cyclopean stimulus and compared with the identical luminance-domain display. The perception of element versus group movement was similar for both stimulus domains. Element movement predominated at short interstimulus intervals, while group movement predominated at long interstimulus intervals. The classic interpretation of bistable motion is that element movement relies on a lower-level, short-range motion system, whereas group movement relies on a higher-level, long-range system. This interpretation is challenged by the cyclopean results where both element and group percepts are supported by a high-level, second-order stimulus.

6. PATTERSON R, RICKER C, MCGARY J, ROSE D: **Properties of Cyclopean Motion Perception.** *Vision Res* 1992, 32:149–156.

7. PANTLE A: **Immobility of Some Second-Order Stimuli in Human Peripheral Vision.** *J Opt Soc Am [A]* 1992, 9:863–867.

8. MATHER G: **First-Order and Second-Order Visual Processes in the Perception of Motion and Tilt.** *Vision Res* 1991, 31:161–167.

9. • WERKHOVEN P, SPERLING G, CHUBB C: **The Dimensionality of Texture-Defined Motion: a Single Channel Theory.** *Vision Res* 1993, in press.

An ambiguous apparent motion display is used in which one potential motion path matches two different texture patches, while the competing path matches two similar patches. In some circumstances, the authors found that motion between patches of different texture could be stronger than motion between patches of similar texture. The results support a univariant model of motion-from-texture in which motion strength is computed from a single spatial transformation of the stimulus, modeled as the rectified output of a low-pass spatial filter applied to stimulus contrast.

10. • CAVANAGH P, ANSTIS SM: **The Contribution of Color to Motion in Normal and Color-Deficient Observers.** *Vision Res* 1991, 31:2109–2148.

Superimposed luminance and color gratings were set in motion in opposite directions and adjusted to produce a motion null between the two gratings. The motion response for a low spatial and low temporal frequency red/green grating was about half as effective as that from a luminance grating where both stimuli were considered in terms of the cone contrast they produced. On the other hand, there was only a very modest input, if any, to the motion pathway from the short-wavelength sensitive cones (B-cones).

11. CAVANAGH P, MATHER G: **Motion: the Long and Short of it.** *Spatial Vision* 1989, 4:103–129.

12. • TURANO K: **Evidence for a Common Motion Mechanism of Luminance-Modulated and Contrast-Modulated Patterns: Selective Adaptation.** *Perception* 1991, 20:455–466.

Adaptation effects were measured with contrast-modulated gratings (a second-order stimulus) and luminance-modulated gratings. Contrast-modulation thresholds for a drifting contrast-modulated test were raised significantly following adaptation to a luminance-modulated grating moving in the same direction as the test, relative to thresholds obtained following adaptation to a luminance-modulated grating moving in the opposite direction. This direction specific adaptation was also found for the reverse pairing: contrast-modulated adapting stimulus and a luminance-modulated test.

13. STONER GR, ALBRIGHT TD: **Motion Coherency Rules are Form-Cue Invariant.** *Vision Res* 1992, 32:465–475.

14. ADELSON EH, MOVSHON JA: **Phenomenal Coherence of Moving Visual Patterns.** *Nature* 1982, 300:523–525.

15. DERRINGTON AM, BADCOCK DR: **Two-Stage Analysis of the Motion of 2-Dimensional Patterns: What is the First Stage?** *Vision Res* 1992, 32:691–698.

16. • DERRINGTON AM, BADCOCK DR, HOLROYD SA: **Analysis of the Motion of 2-Dimensional Patterns: Evidence for a Second-Order Process.** *Vision Res* 1992, 32:699–707.

Two gratings of the same spatial frequency and contrast were superimposed to generate a plaid pattern. The gratings moved in discrete jumps. When the size of the jumps was increased to 3/8 of the gratings' spatial period, the perceived direction of motion of the plaid pattern reversed, even though neither component reversed when viewed in isolation. Results suggest that the plaid creates a new grating of twice the spatial frequency with alternating bars of different texture, either uniform or striped. Responses of a second-order motion mechanism to this emergent texture grating contribute to the analysis of plaid motion.

17. YO C, WILSON HR: **Perceived Direction of Moving Two-dimensional Patterns Depends on Duration, Contrast and Eccentricity.** *Vision Res* 1992, 32:135–147.

18. STELMACH LB, HERDMAN CM: **Directed Attention and Perception of Temporal Order.** *J Exp Psychol [Hum Percept Perform]* 1991, 17:539–550.

19. •• HIKOSAKA O, MIYAUCHI S, SHIMOJO S: **Focal Visual Attention Produces Illusory Temporal Order and Motion Sensation.** *Vision Res* 1993, in press.

The spatial distribution of attention is often evaluated by measuring its facilitating effect on reaction time, percent correct detection, and the speed of perception of events [18]. In this paper, a new technique is shown to have very high sensitivity. When a probe line is flashed with one end in an attended region, a distinctive motion is seen along the length of the line. The experiments show that this effect can be reliably quantified and even nulled by turning on the line incrementally from the other direction.

20. •• CAVANAGH P: **Attention-Based Motion Perception.** *Science* 1992, 257:1563–1565.

Two 'attentive' tracking tasks reveal the existence of an attention-based motion process. In the first task, oppositely rotating luminance and color gratings were superimposed. Because of masking from the color grating, the bars of the luminance grating were not visible; nevertheless, their motion was visible and it determined the perceived direction of the stimulus rotation. On the other hand, the bars of the color grating were visible, but they could only be seen to move (in the opposite direction to the overall stimulus rotation) when they were tracked with attention. In a second task, the perceived velocity of a color grating, typically slow at equiluminance, increased when individual bars were attentively tracked.

21. PYLYSHYN ZW, STORM RW: **Tracking Multiple Independent Targets: Evidence for a Parallel Tracking Mechanism.** *Spatial Vision* 1988, 3:151–224.

22. YANTIS S: **Multi-Element Visual Tracking: Attention and Perceptual Organization.** *Cogn Psychol* 1992, 24:295–340.

23. •• KAHNEMAN D, TREISMAN A, GIBBS BJ: **The Reviewing of Object Files: Object-Specific Integration of Information.** *Cogn Psychol* 1992, 24:175–219.

In this report the authors demonstrate several object-specific priming effects. In general, a preview field containing two or more letters was followed by a target letter to be named, and then the response time was measured. Displays were designed to induce the integration of prime and target as successive states of a single object. The integration of the two states was suggested in the different experiments by either a shared location, a shared relative position in a moving pattern, or successive appearance in the same moving frame.

24. •• AHUMADA AJ JR: **Learning Receptor Positions.** In *Computational Models of Visual Processing.* Edited by Landy MS, Movshon JA. Cambridge, Massachusetts: MIT Press; 1991: 23–34.

A neural network model is examined for a visual system whose photoreceptors are initially at unknown locations. Two learning procedures are presented that converge on a regularized internal representation. The methods can also compensate for optical distortions such as spherical aberration. They work by comparing visual input across eye/head movements and require no explicit feedback and no knowledge about the particular contents of a scene.

25. RAMACHANDRAN VS, BRADDICK O: **Orientation-Specific Learning in Stereopsis.** *Perception* 1973, 2:371–376.

26. MCKEE SP, WESTHEIMER G: **Improvement in Vernier Acuity with Practice.** *Percept Psychophys* 1978, 24:258–262.

27. FIORENTINI A, BERARDI N: **Perceptual Learning Specific for Orientation and Spatial Frequency.** *Nature* 1980, 287:43–44.

28. Ball K, Sekuler R: **A Specific and Enduring Improvement in Visual Motion Discrimination.** *Science* 1982, 218:697–698.

29. Karni A, Sagi D: **Where Practice Makes Perfect in Texture Discrimination: Evidence for Primary Visual Cortex Plasticity.** *Proc Natl Acad Sci USA* 1991, 88:4966–4970.
••
Subjects reported the orientation of a small rectangular patch of oblique line segments within a background of horizontal line elements. Threshold exposure durations were evaluated and the task was practiced for many days. Very pronounced practice effects were found to be local (in a retinotopic sense), orientation specific, but asymmetric (specific for background but not for target-element orientation), and strongly monocular (there was little interocular transfer of learning). These results suggest local plasticity at the level of orientation-gradient sensitive cells in primary visual cortex.

30. Shiu L-P, Pashler H: **Improvement in Line Orientation Discrimination is Retinally Local but Dependent on Cognitive Set.** *Percept Psychophys* 1992, 52:582–588.

31. Poggio T, Fahle M, Edelman S: **Fast Perceptual Learning in Visual Hyperacuity.** *Science* 1992, 256:1018–1021.

32. Nakayama K, Shimojo S: **Experiencing and Perceiving Visual Surfaces.** *Science* 1992, 257:1357–1363.

33. Cavanagh P: **What's up in Top-Down Processing?** In *Representations of Vision: Trends and Tacit Assumptions in Vision Research*. Edited by Gorea A. Cambridge, UK: Cambridge University Press; 1991: 295–304.

34. Poggio T, Edelman S: **A Network that Learns to Recognize Three-Dimensional Objects.** *Nature* 1989, 343:263–266.

35. Bulthoff HH, Edelman S: **Psychological Support for a Two-Dimensional Interpolation Theory of Object Recognition.** *Proc Natl Acad Sci USA* 1992, 89:60–64.
•
Subjects were trained to discriminate wire-frame or amoeba-like 3D shapes presented at two training orientations. These figures were rotated slightly around these orientations to provide a reliable 3D impression of their shape, and in one condition stereo information was provided as well. Recognition of the shapes was then tested at various orientations between the training values or outside that range. All results indicated poor recognition of radically unfamiliar views. Even when given every opportunity to generate an internal 3D representation for future matching, the subjects performed as if they had not done so. The results suggest that under a variety of conditions recognition is based on stored 2D views, and involves neither 3D object models nor explicit compensation for viewpoint variability.

P Cavanagh, Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, Massachusetts 02138, USA.