MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY
and
CENTER FOR BIOLOGICAL AND COMPUTATIONAL LEARNING
DEPARTMENT OF BRAIN AND COGNITIVE SCIENCES

A.I. Memo No. 1506
C.B.C.L. Memo No. 104
November, 1994

Reciprocal interactions between motion and form perception

Pawan Sinha
This publication can be retrieved by anonymous ftp to publications.ai.mit.edu.

Abstract

The processes underlying the perceptual analysis of visual form are believed to have little interaction with those subserving the perception of visual motion [Livingstone and Hubel, 1987; Victor and Conte, 1990]. Recent reports of functionally and anatomically segregated parallel streams in the primate visual cortex seem to support this hypothesis [Ungerlieder and Mishkin, 1982; VanEssen and Maunsell, 1983; Shipp and Zeki, 1985; Zeki and Shipp, 1988; De Yoe et al., 1994]. Here we present perceptual evidence that is at odds with this view and instead suggests strong symmetric interactions between the form and motion processes. In one direction, we show that the introduction of specific static figural elements, say 'F', in a simple motion sequence biases an observer to perceive a particular motion field, say 'M'. In the reverse direction, the imposition of the same motion field 'M' on the original sequence leads the observer to perceive illusory static figural elements 'F'. A specific physiological implication of this finding concerns the possible existence of (what we call) motion end-stopped units in the primate visual system. Such units might constitute part of a mechanism for signalling subjective occluding contours based on motion-field discontinuities.

This report describes research done at the Artificial Intelligence Laboratory and the Center for Biological and Computational Learning of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-91-J-4038. The Center is supported in part by NSF-ASC-9217041 and by a grant from the National Institutes for Health under contract NIH 2-S07-RR07047. The author can be reached at: NE43-809, MIT AI Laboratory, 545 Technology Square, Cambridge, MA 02139, USA. E-mail: sinha@ai.mit.edu

1 Introduction

Our claim of strong reciprocal interactions between the analyses of two-dimensional visual form and motion rests on the demonstration that for a large class of simple motion sequences, changes in the perceived static 2D figural elements can cause, and be caused by, modifications in the perceived 2D motion-field. The motion sequences we use are two-dimensional animations comprising of simple geometric figures translating under stationary occluding patches. These patches obscure the moving figures' vertices. Figure 1(b) shows a typical sequence (for ease of exposition, we shall work exclusively with this sequence (call it sequence 'A') in this paper). It was created by moving a diamond shaped wireframe horizontally back and forth under three stationary vertical occluders (figure 1(a)). The diamond's extent of translation was limited to ensure that none of its vertices were ever visible. The occluding strips were assigned exactly the same gray-level as the background, thereby rendering their boundaries invisible.

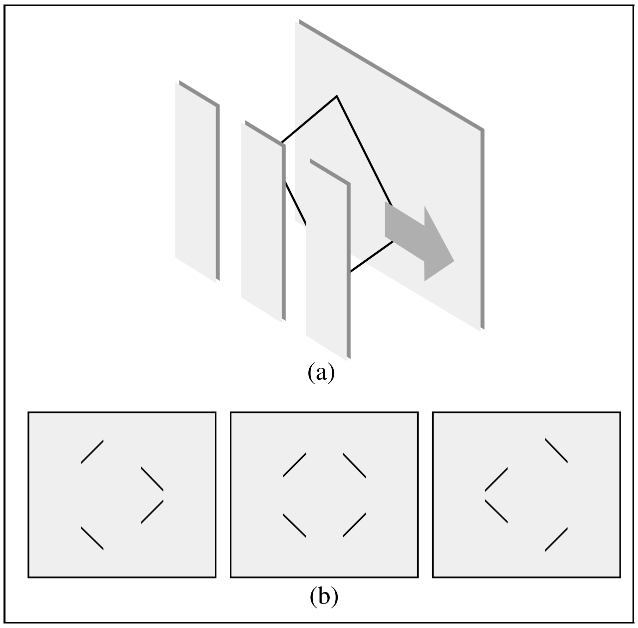

Figure 1: *(a) The setup used in our experiments. (There is no physical significance to the absolute depth differences between the occluders, diamond and background shown in this figure. They are indicative only of the depth ordering used to generate the 2D motion sequences.) We used the Macromind Director software package running on a Macintosh Quadra 700 to generate the sequences. The sequences were presented on a 16" monochrome monitor and subtended 7 degrees of visual angle at a viewing distance of 80 cm. The diamond's speed of oscillation was 0.7 degrees per second. (b) Three frames from sequence 'A': the diamond oscillating horizontally behind occluders that have exactly the same gray-level as the background.*

We performed two sets of experiments. The first set was designed to study how a change in the figural aspects of sequence 'A' could effect its motion percept. The second set examined the possibility of a reciprocal influence - a change in sequence 'A's perceived motion field influencing its perceived figural content.

2 Experiment 1: The influence of form on motion

2.1 Methods:

Our subjects were five MIT graduate students, naive as to the purpose of the experiment. The experiment had two parts. In part one, the subjects were individually shown ten presentations (each lasting 5 seconds) of sequence 'A'. They were instructed to verbally report, after each presentation, the dominant direction of perceived motion (horizontal, diagonal or vertical).

For part two of the experiment, we created a new sequence (sequence 'B') by modifying sequence 'A' to explicitly suggest the presence of the occluding strips (figure2(a)). The presentation schedule for part one was repeated with sequence 'B'.

2.2 Results and Discussion:

Figure 2(b) summarizes the different results obtained with sequences 'A' and 'B'. The dominant percept for all subjects with sequence 'A' was that of vertical oscillatory motion (figure 3(a)) while sequence 'B' was perceived as having motion in the horizontal direction (figure 3(b)). It is important to note that the raw motion signals in sequences 'A' and 'B' are identical; the changes in perceived motion can therefore be justifiably attributed simply to the differences in the static figural contents of the two sequences.

Precisely how might the visibility of the strips in sequence 'B' influence the analysis of motion? There are at least two possibilities: First, with the strips visible, the line terminations are interpreted as arising out of occlusion rather than as intrinsic features of the moving lines. This induces the visual system to discard their vertical motion [Shimojo et al., 1989; Shiffrar and Pavel, 1991; Lorenceau and Shiffrar, 1992]. The only intrinsic motion signals now available are those from the interiors of the line segments. However, these motion signals are inherently ambiguous due to the aperture problem [Hildreth, 1984; Poggio et al. 1985]: for each line segment, only the component of motion perpendicular to its orientation can be recovered. Second, by suggesting the possibility of the line-segments continuing behind them, the strips allow amodal grouping of the segments and the consequent integration of their motion signals, possibly by the intersection of constraints construction [Movshon et al. 1985]. The resultant motion is in the horizontal direction.

The net effect is that instead of seeing four independent line-segments moving vertically, the observer now perceives them as moving coherently in the horizontal direction. A change in the figural content of sequence 'A' thus leads to a change in the perceived motion (figure 4(a)).

The preceding demonstration of the influence of form on perceived motion is not entirely new. Somewhat similar effects have been reported earlier, though, to the best of our knowledge, this is the first demonstration of subjective static figural elements exerting an influence on perceived motion. It is also important to note that this influence is non-local.

The more significant factor that distinguishes this study from the earlier ones is our emphasis on reciprocality; we investigate the possibility of interactions in the opposite direction: can a modification in perceived motion influence perceived form? Or, in the current context, can the 'imposition' of a horizontal motion field on sequence 'A' result in the perception of the form of the occluding strips (figure 4(b))? Experiment 2 attempts to address this question.

3 Experiment 2: The influence of motion on form

3.1 Methods:

Our subjects for this experiment were five naive MIT graduate students, different from the ones who had participated in experiment 1. This was done to prevent an influence of any experientially acquired biases from experiment 1 on the subjects' performance. The experiment had two parts. In part one, the subjects were individually shown ten presentations (each lasting 5 seconds) of sequence 'A'. They were instructed to indicate with a cursor on the display, after each presentation, the locations of any perceived subjective contours.

For part two of the experiment, we created a new sequence (sequence 'C') by 'imposing' a horizontal motion field on sequence 'A'. This was done by attaching to the diamond's contour point-features moving unambiguously in the horizontal direction (figures 5(a) and 5(b)). The point features were so placed as to never be obscured by occluders. The presentation schedule for part one was repeated with sequence 'C'.

3.2 Results and Discussion:

Figure 6(a) summarizes the different results obtained with sequences 'A' and 'C'. The subjects reported perceiving strong illusory contours [Petry and Meyer, 1987] in nearly all presentations of sequence 'C' but not in sequence 'A'. The responses of all subjects were in agreement regarding the locations of the perceived contours. The contours were seen as delimiting illusory occluding strips at the locations indicated in figure 6(b). Any single static frame from this sequence did not afford the percept of subjective contours. Interestingly, on subsequent presentations of sequence 'A', four of the subjects reported being able to perceive subjective occluding boundaries simply by mentally imposing a horizontal motion field on the sequence; the subjective contours disappeared when the four line segments were seen as moving vertically. In essence, then, changes in the motion field, whether imposed externally or internally via top-down mechanisms, profoundly influenced the perceived figural content for our subjects.

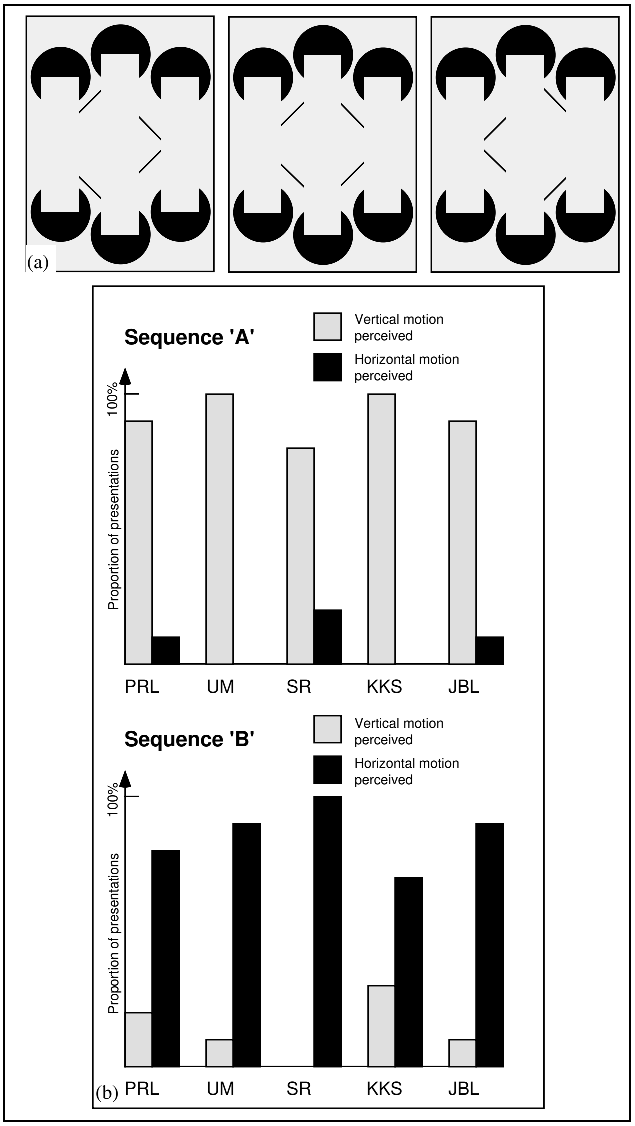

Figure 2: *(a) Three frames from sequence 'B'. This is the same as sequence 'A' except for the inclusion of a few static figural elements that indicate the occluders' presence. (b) A summary of the results obtained with sequences 'A' and 'B' in experiment 1. For each of the five subjects, the inclusion of the static figural elements strongly biases the percept from being one of motion in the vertical direction to one of horizontal motion.*

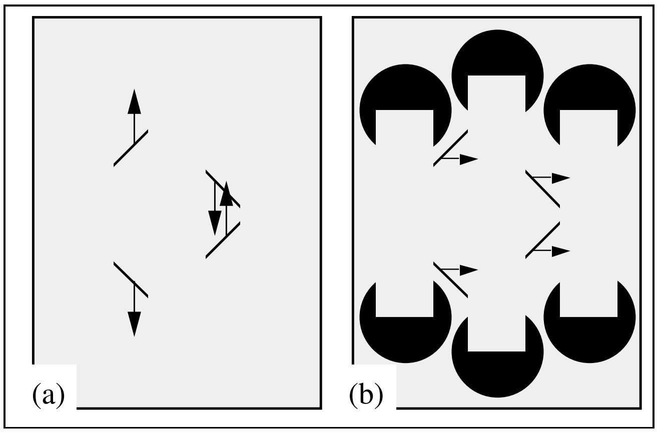

Figure 3: *(a) The reported percept with sequence 'A': four lines oscillating vertically independently of each other (b) The reported percept with sequence 'B': a partially obscured diamond translating horizontally back and forth behind three opaque strips.*

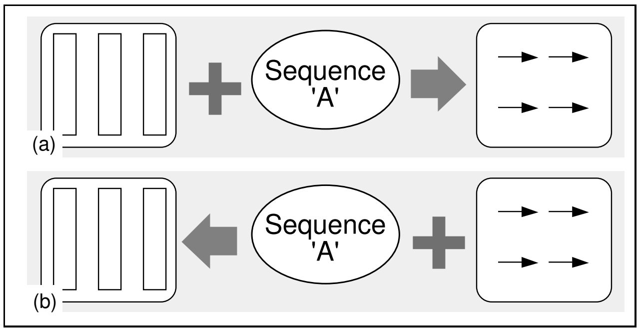

Figure 4: *(a) and (b) The paradigm to test for reciprocal interactions between the perception of 2D form and motion. Panel (a) schematically summarizes the result obtained with experiment 1: introducing the explicitly indicated occluding strips in sequence 'A' yields a horizontal motion-field. Panel (b) is the test for reciprocity: would the imposition of a horizontal motion-field on sequence 'A' yield a percept of the occluding strips? Experiment 2 addresses this question.*

This demonstration of motion influencing the perception of contours makes intuitive sense when one notices that the locations of the subjective contours coincide with the discontinuities in the horizontal motion-field (figure 6(b)). Given that occluding surfaces often result in motion discontinuities, it seems reasonable for the visual system to use the latter to hypothesize the presence of the former (Michotte invoked a similar idea in some of his work on the perception of causality [Michotte, 1954]). Furthermore, it is only meaningful to infer occlusion from discontinuities along the motion trajectory rather than those perpendicular to it. This is consistent with our aforementioned result that subjective contours are perceived when the four segments are seen as moving together horizontally, but not when they are seen as moving vertically.

It is worth emphasizing that our use of the term 'motion discontinuities' has a somewhat different meaning than that implied in conventional usage. The motion discontinuities in our sequences are punctate - at any time, discontinuities exist at only a few points. This is in contrast to more conventional settings where motion discontinuities are used to refer to spatially extended and continuous contours that demarcate two different motion fields. The remarkable feature about the percept obtained with sequence 'C' is that it demonstrates the visual system s ability to infer the presence of extended occluding contours from punctate motion discontinuities. This is completely analogous to the situation in the contrast domain where given only a few punctate contrast discontinuities in a Kanizsa-like figure, the visual system infers extended subjective contours.

4 Physiological implications:

The perceptual results we have described provide some hints regarding the nature of the corresponding physiological processes. We mention a few speculations below.

Experiment 1 suggests that the percept of 2D pattern motion can be modified by varying the figural content of a motion sequence. Given that the medial temporal (MT) area is thought to be intimately involved in the computation of pattern motion [Movshon et al., 1985; Albright, 1984; Newsome and Pare, 1988; Rodman and Albright, 1989], this implies that it should be possible to modulate the activity of at least some of the MT units by simply modifying some static figural characteristics of the presented sequence. Furthermore, these modifications can be non-local, i.e. they need not lie within the receptive field of the unit being monitored. Support for this idea comes from the fact that the results in experiment 1 were obtained by the inclusion of static forms (partially occluded disks) that left the local patterns of motion physically unchanged. These findings suggest a conceptually simple physiological test. The test would involve presenting sequences 'A' and 'B' within the receptive field of an MT cell tuned to motion in the horizontal direction. The interesting question to ask then would be: does the cell respond more with sequence 'B'

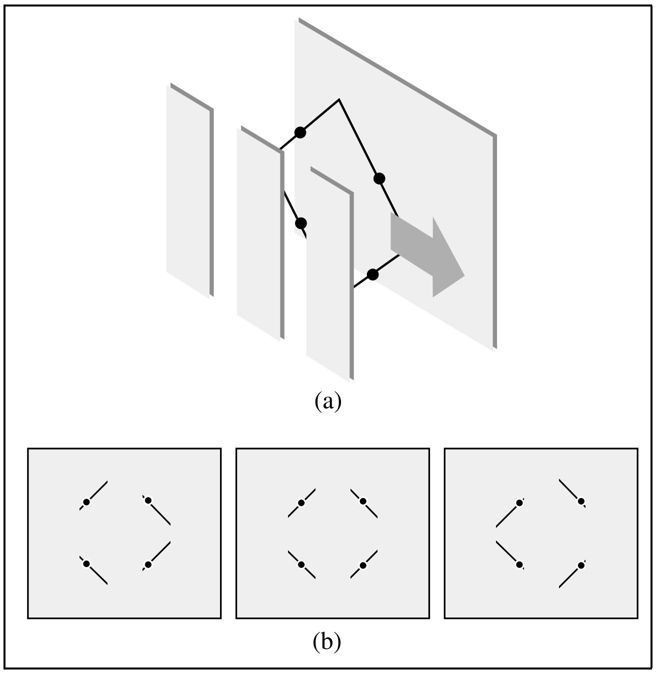

Figure 5: *(a) The setup used to generate sequence 'C'. The moving diamond now has four point-features that render it immune to the aperture problem: its horizontal motion is now clearly evident. (b) Three frames from sequence 'C'.*

than with sequence 'A' even though the physical motion content in the two is identical? In other words, does the presence or absence of the partially occluded disks modulate the response of an MT cell? If such a modulation were to be found, an obvious question would concern the source of such modulating influences. Area V2, which has been implicated in form perception and has been shown to have several cells responsive to subjective contours [Peterhans and von der Heydt, 1989, 1991], would be a likely candidate.

Let us now consider experiment 2. The results with experiment 2 imply that there must exist a physiological mechanism capable of signalling occluding contours based on motion field discontinuities. Given that this mechanism would use motion signals as input and yield form related output, it would be appropriate to think of it as part of a larger scheme for interaction between processing in the two domains. Building upon the body of work exploring the use of contrast discontinuities to signal occluding contours [Peterhans and von der Heydt, 1989, 1991; Heitger et al., 1992, 1993; Dobbins et al. 1989], we suggest that a key component of such a mechansim might be (what we call) a 'motion end-stopped' unit. Such a unit would be well suited to detecting motion-field discontinuities along the motion trajectory (figure 7).

We need to be a little more specific about the nature of the motion-field that these units might operate on. The motion-field can either comprise of local mo-

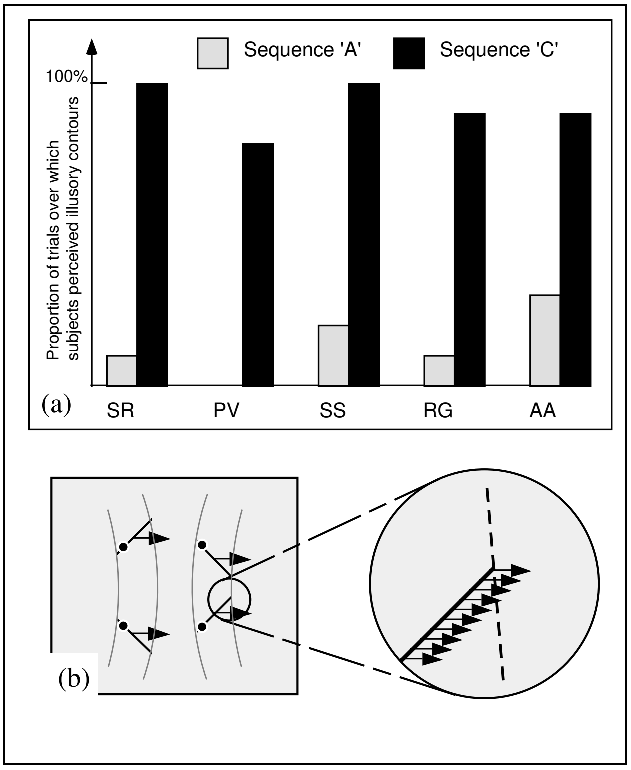

Figure 6: *(a) A summary of the results obtained with experiment 2. The imposition of a horizontal motion field induced subjects to perceive contours delimiting illusory occluding strips. (b) The percept with sequence 'C': strong subjective contours are reported at the indicated locations. As shown in the enlarged inset, the illusory contours coincide with the abrupt discontinuities in the horizontal motion-field. The reported contours were not precisely vertical but slightly curved, as indicated. We have yet to systematically study the cause for this perceived curvature.*

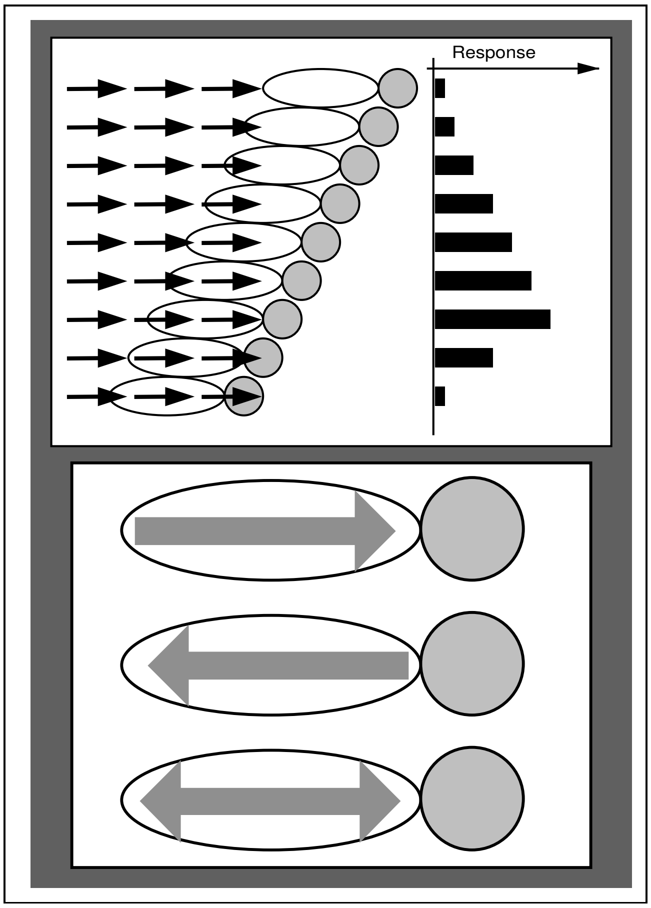

Figure 7: *A qualitative characterization of the predicted 'motion end-stopped' units. The top panel shows how the response is expected to vary as a function of the extent of overlap between a motion field and a unit's receptive field. The excitatory and inhibitory regions of the receptive field are represented by the white and speckled ovals respectively. As indicated in the lower panel, the preferred direction of the unit needs to be aligned with its major axis to allow it to detect discontinuities in a motion-field along the motion trajectory. The discontinuities can be classified into two basic categories: those associated with the abrupt appearance of motion and those associated with its abrupt disappearance. These discontinuities can either be detected by two different units with opposite preferred directions or by one unit that is sensitive to the motion field's orientation but not its sign.*

tion measurements (such as those available at the level of the striate cortex) or of motion signals obtained by globally integrating the local measurements (such integration could be performed by the medial temporal area). We investigate this issue using a sequence ('D') that yields motion fields that are locally identical to but globally different from those produced by sequence 'C'. Sequence 'D' is the same as sequence 'C' except for having the diamond oscillate vertically instead of horizontally (figure 8(a)). Owing to the aperture problem, a vertically oscillating diamond yields local patterns of motion that are identical to those produced by a horizontally moving one (figure 8(b)). (More specifically, over the course of every oscillation cycle, sequence 'D' yields the same local patterns of motion of the eight line terminations as sequence 'C', albeit with different phase shifts.) If local motion measurements subserved the perception of contours, sequence 'D' would be expected to yield the percept of occluding boundaries just as sequence 'C' does. On the other hand, if the contour signalling mechanism operated on the global motion field, the prediction would be very different. Sequence 'D' has no discontinuities along the motion trajectories in its global motion field (figure 8(c)) and would not, therefore, be expected to yield the percept of contours. The results unequivocally support the latter alternative. None of the five subjects tested (the participants of experiment 2) reported perceiving any subjective contours in sequence 'D'.

Taken together, our findings with sequences A, B and C strongly suggest tight reciprocal interactions between processing mechanisms in the 2D form and motion domains. We have obtained similar results with several other such sequences in both the 2D and 3D domains. We are therefore led to suggest that to be able to account for these results, the notion of independent streams of analysis in the visual system needs to be augmented with mechanisms to permit extensive crosstalk. Our hypothetical 'motion end-stopped' unit might be part of one such mechanism.

Acknowledgements:

The author wishes to thank Dr. Tomaso Poggio and Dr. Heinrich Buelthoff for their helpful comments.

References:

Albright, T. D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. Journal of Neurophysiology, 52, 1106-1130.

De Yoe, E. A., Felleman, D. J., Van Essen, D. C., McClendon, E. (1994). Multiple processing streams in occipitotemporal visual cortex. Nature, 371(8): 151-154.

Dobbins, A., Zucker, S. W. and Cynader, M. S. (1989). Endstopping and Curvature. Vision Research. 10, 1371-1387.

Heitger, F., Rosenthaler, L., von der Heydt, R., Petehans, E., and Kubler, O. (1992). Simulation of neural contour mechanisms: from simple to end-stopped cells.

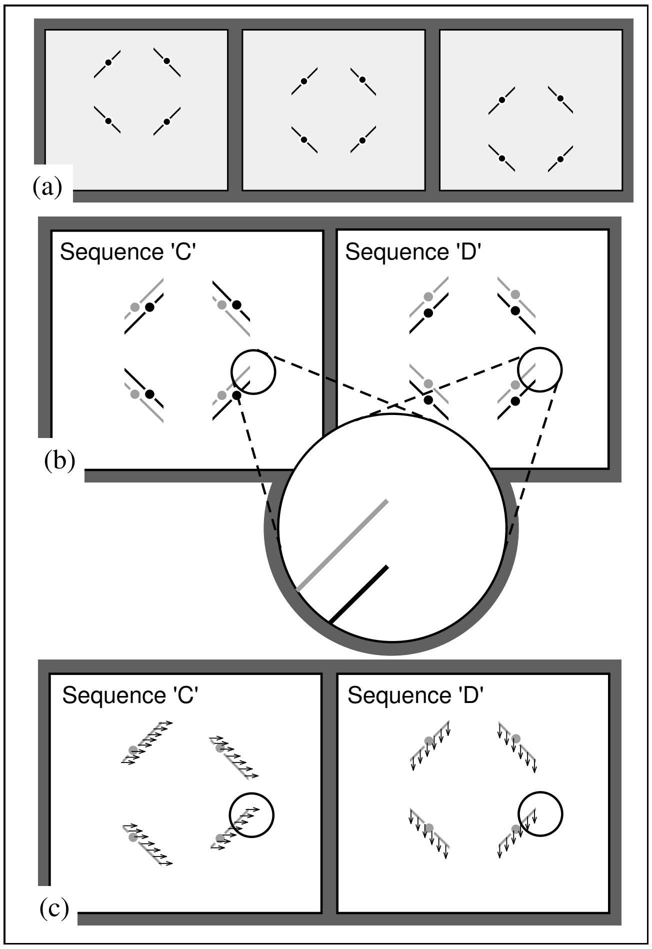

Figure 8: *(a) Three frames from sequence 'D'. (b) Due to the aperture problem, the local patterns of motion in sequences 'C' and 'D' are identical (they do, however, differ in phase). (c) The global motion fields in the two sequences are, however, very different.*

Vision Research. 32(5): 963-981.

Heitger, F. and von der Heydt, R. (1993). A computational model of neural contour processing: Figure-ground segregation and illusory contours. In Proceedings of the ICCV, 32-40, Berlin, Germany.

Hildreth, E. (1984). The measurement of visual motion. Cambridge, MA: MIT Press.

Livingstone, M. S. and Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement and depth. Journal of Neuroscience, 7, 3416-3468.

Lorenceau, J. and Shiffrar, M. (1992). The influence of terminators on motion integration across space. Vision Research, 32: 263-273.

Michotte, A. (1954). La Perception de la Causalite'. Louvain: Publications Universitaires.

Movshon, J. A., Adelson, E. H., Gizzi, M. S. and Newsome, W. T. (1985). The analysis of moving visual patterns. In Pattern Recognition Mechanisms, Eds. C. Chagas, R. Gattass, and C. Gross, 117-151. Vatican City: Pontifical Academy of Science.

Newsome, W. T. and Pare, E. B. (1988) A selective impairment of motion perception following lesions of the middle temporal visual area (MT). J. Neuroscience, 8: 2201-2211.

Peterhans, E., and von der Heydt, R. (1989a) Elements of form perception in monkey V2 - A correlation with the cytochrome oxidase pattern. Soc. Neurosci. Abstr. 15: 161.

Peterhans, E., and von der Heydt, R. (1991) Elements of form perception in monkey prestriate cortex. In Representations of Vision. Eds A. Gorea, Y. Fregnac, Z. Kapoula, J. Findlay, pp.111-124. Cambridge: Cambridge University Press.

Petry, S., and Meyer, G. L. (1987) The perception of illusory contours. New York: Springer.

Poggio, T., Torre, V., and Koch, C. (1985). Computational vision and regularization theory. Nature, 317, 315-319.

Rodman, H. R. and Albright, T. D. (1989). Single unit analysis of pattern motion selective properties in the middle temporal visual area MT. Experimental Brain Research, 75, 53-64.

Shiffrar, M. and Pavel, M. (1991). Percepts of Rigid motion within and across apertures. Journal of Expt. Psychology: Human Perception and Performance, Vol. 17, No. 3, 749-761.

Shimojo, S., Silverman, G. and Nakayama, K. (1989). Occlusion and the solution to the aperture problem for motion. Vision Research, 29(5): 619-626.

Shipp, S. D. and Zeki, S. (1985). Segregation of pathways leading from area V2 to areas V4 and V5 of macaque monkey visual cortex. Nature, London, 315, 322-324.

Ungerlieder, L. G. and Mishkin, M. (1982). Two cortical visual systems. In: The analysis of visual behavior, edited by D. J. Ingle, M. A. Goodale, and R. J. W. Mansfield. Cambridge, MA; MIT Press, 549-586.

Van Essen, D. C. and Maunsell, J. H. R. (1983). Hierarchical organization and functional streams in the visual cortex. Trends in Neuroscience, 6, 370-375.

Victor, J. D. and Conte, M. M. (1990). Motion mechanisms have only limited access to form information. Vision Research, Vol. 30, No. 2, 289-301.

Zeki, S., and Shipp, S. (1988). The functional logic of cortial connections. Nature (London) 335: 311-317.