Can Early Stage Vision Detect Topology

Liu, Lifu Zhao, Nanyuan Bian, Zhaoqi
Department of Automation, Tsinghua University
Beijng 100084, P. R. China

Abstract

The apparent motion reveals what in an image that human vision detects first. Chen's assumption that early stage vision can percept global topology is proved incorrect in theory and experiments. Based upon psychological studies on human vision, a new theory, the blurred matching, was introduced into visual computation which well fits the results of all experiments about apparent motion, Chen's included. An algorithm was constructed accordingly which predicts apparent motion by comparison of distances.

1 Introduction

What is the mechanism of human visual perception, and how does it function in the process of "seeing"? What are detected first in the process? The answer to the questions determines how our visual perception theory should be established.

Marr of MIT, for instance, gave his answer in which he assumed that visual perception represents a symbolic description of the outside world, and the understanding towards the world proceeds as computation on these symbolic representations or elements, among them the most primitive include visual information like variation of brightness and local geometry (parallelism, relative position and orientation). Any image of the surrounding world is interpreted through computation on the elements. Based upon the assumption, Marr established his theory of vision: visual perception is computation.

However, L. Chen questioned Marr's idea. He claims that one essential geometrical property is excluded in Marr's assumption of the most primitive elements of visual information, the topology. The computational complexity of topology is so high, according to Minsky's theorem, that no topological invariance except the Eular number E(x) may be discriminated by a finite order perception machine. In other words, such a machine is unable to resolve topological properties like connectedness and closedness. Since the perception machine may also be viewed as a parallel Turing machine, the detection of topology in early stage vision reveals a fatal defect of Marr's theory as well as vision computation in Turing sense, because the problem is unconfutable by any mathematical means currently available

The objective of the paper, after reviewing related research background, is to make clear whether early stage vision can detect topology, a question concerns whether vision is computer simulatable Its significance to vision research can not be overestimated for it shows if the work itself is worthwhile.

Starting from Chen's conclusion, we are going to point out in the following passages his error when he denounced Ramachandran's low spacial frequency theory to set his own of detection of topology, and how he misinterpreted the results of his experiment. In fact, his conclusion is false. This is demonstrated by a number of experiments of apparent motion, which deny topology as one of the most primitive elements of visual information. We will then introduce the concept of blurred matching to explain phenomena of apparent motion, and provide a practical algorithm for prediction of moving direction.

2 Notes on Experiments

One most commonly used method to study early stage vision is through experiments of apparent motion on tachistoscope. With the images Chen used in his experiments, we briefly describe how the experiment is done.

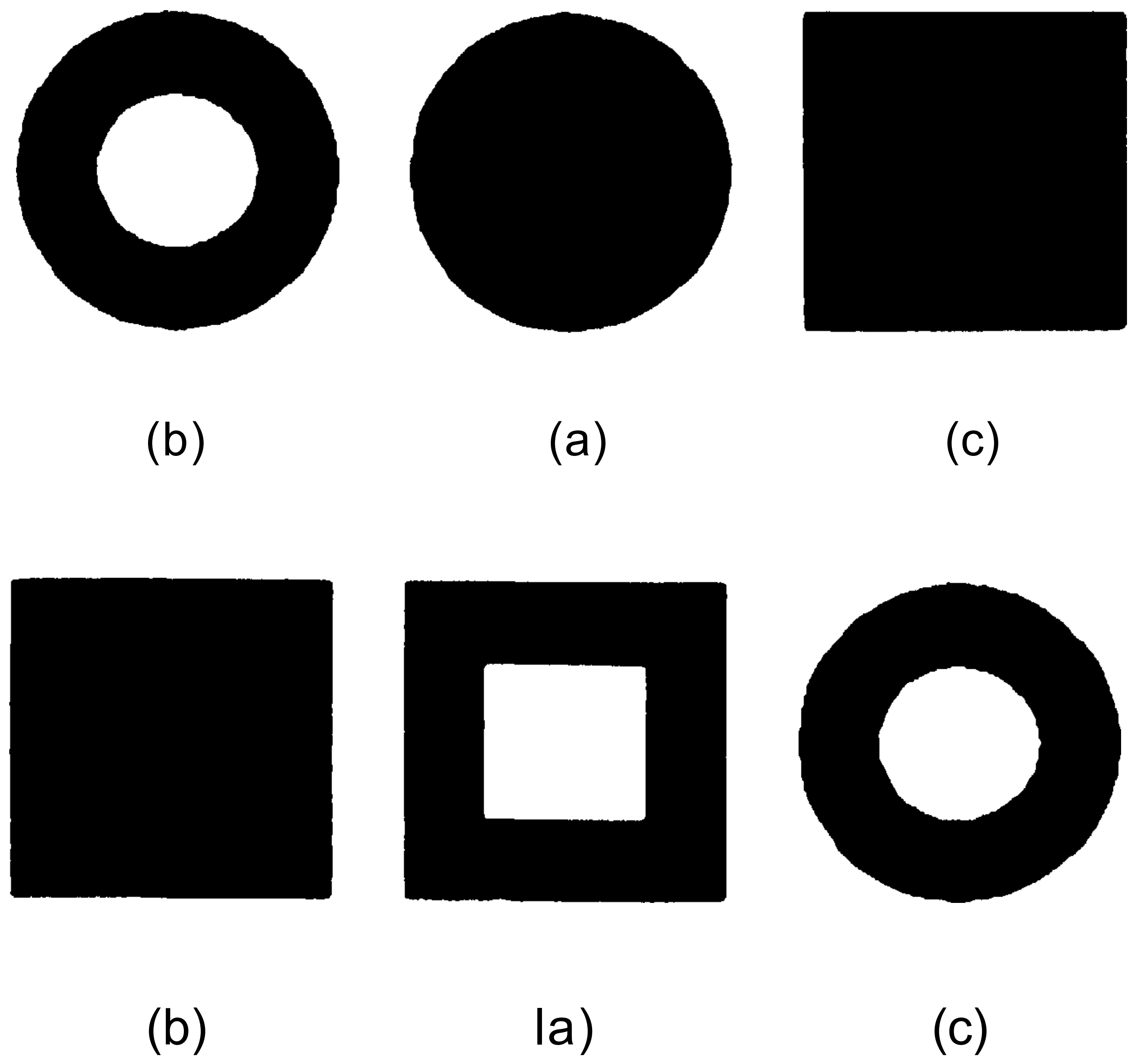

Fig. 1 Images used by Chen. More subjects reported apparent motion of (a) to (c), which are topologically equivalent.

In each experiment, two stimulative images are displayed asynchronously through tachistoscope. The first have a single centered stimulative pattern(see pattern a in Fig 1), and the second contains two patterns equally spaced from the center of the image(patterns b and c). Both images have their own display time and there is a time interval between the two successive displays.The subjects under such display condition will have a feeling of pattern (a) moving towards either pattern (b) or (c), and he is asked to report the moving direction - leftwards or rightwards. In Chen's experiments, more subjects reported apparent motion of pattern (a) to *(c.)* in Fig. 1. This, he attributed, is because of (a) and (c) are topologically equivalent.

3 Misinterpretation to Ramachandran's Theory

Ramachandran pointed out that low spacial frequency area plays a decisive role in perception of apparent motion. It is necessary for Chen to rule out influences by all other factors, including the low spacial frequency, before he could justify the key role that topology plays. This is how he does: Those who believing in the Ramachandran will interpret the tendency of apparent motion of bold circles towards bold squares(Fig. 1) as they both contain high percentage of low spacial frequency. But to keep in consistency with the spacial frequency interpretation, the moving tendency of hollow square towards circle in the second triple of patterns in Fig. 1 must be explained as they both contain more high spacial frequency components than bold square does. This high frequency, together with the low frequency interpretation for the first triple, suggests the similarity of shapes determines apparent moving direction, an assumption proved false by many researchers. This compells us to abandon the spacial frequency interpretation and accept that of topology.

Yet Ramacheadran's low spacial frequency areas should refer to areas of identical brightness. This comes from the hypothesis by B. Julesz of AT & T and O. Braddick of the University of Cambrige that visual system may have detected the correspondence of similar low spacial frequency areas before the detection of more detailed contours or apparent edges. We can then see Ramachandran's theory means that our vision tends to match, in comparison to apparent contours, large areas of identical brightness, namely the low spacial frequency areas, regardless of bright or dark. The correct interpretation for the experiment with the first triple of patterns is that (a) and (c) both have a large area of the same brightness - the dark area This rule can be applied to the second triple of patterns in Fig. 1 without cumbersome involvement of high spacial frequency: (a) moves towards (c) because there is no matching bright area in (b). On the other hand, high frequencies corresponds to edges rather than what lies among them, or the low spacial frequency areas.

Thus we showed Chen's proof does not hold because of his misinterpretation of the effects of low spacial frequency

4 No Perception for Topology

To test Chen's theory, we conducted many experiments. Here are some patterns used in them.

The three pattern triples were displayed on terminal of Silicon Graphics workstation IRIS-400. (Programs are available on request). Patterns (a) and (c) in each of the triples are topologically equivalent. In our experiments, (a) always moved towards pattern (b) instead of its topological equivalence (c), in contradictory with what predicted by Chen's topology detecting theory. These contrary examples demonstrated undoubtedly that apparent motion perception can not be explained through topological equivalence.

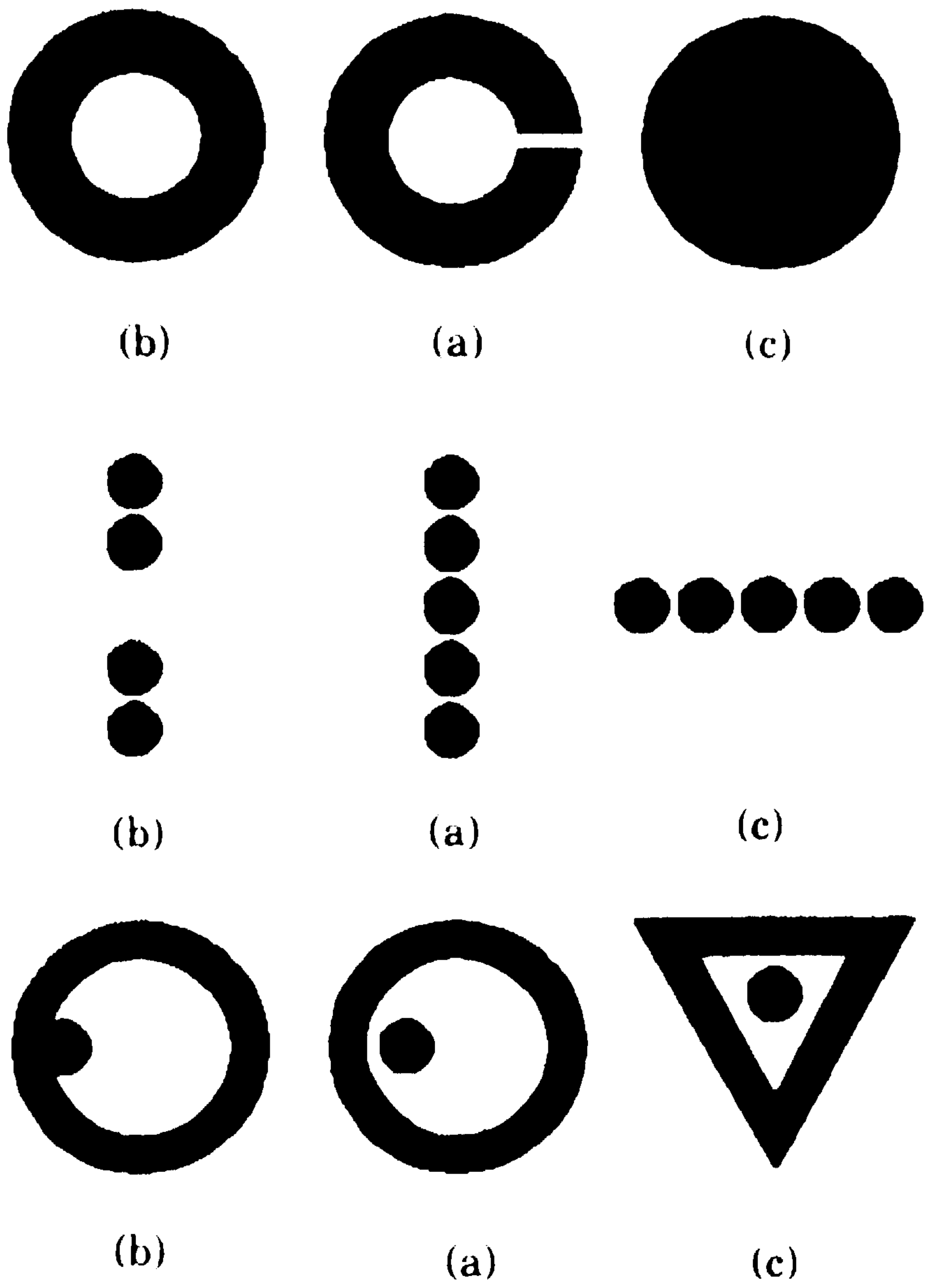

Fig. 2 Pattern (a) always moves to (b), instead of, as predicted by Chen's theory, its topological equivalent (c)

5 Blurred matching

What, then, acts in apparent motion perception? Let's tackle the problem from a new viewpoint.

From physiological research we know that there are two classes of receptors over the surface of the retina: cones and rods. The cones in each eye number between six and seven million. They are located primarily in central portion of the retina, called the fovea. Humans can resolve fine details with these cones largely because each one is connected to its own nerve end. The number of rods is much larger, being in the order of 75 to 120 million distributed over the retinal surface. The larger area of distribution and the fact that several rods are connected to a single nerve end reduce the amount of detail discernible by these receptors. Rods serve to give a general, overall picture of the field of view.

Now let us look into the way a subject percepts images under above experiments. He is asked to fix his eyes at the center of images displayed, where pattern (a) appears Patterns (b) and (c) is blurred to him because they are away from the center and sensed by the rods. It is impossible for him to move his eyes from pattern (a) to (b) or (c), because the display refreshes so rapidly in miliseconds. Suppose (a) is exactly the same as (b), it will certainly moves towards (b). Now let us replace (b) with a somewhat different pattern whose blurred image is just like blurred (b). We should expect the subject still reports motion of (a) to (b), because there is no change to what he percepts The reported direction will not reverse suddenly if shape of (b) changes gradually. Therefore we believe that the moving direction is determined by how close (a) is to blurred (b) or blurred (c). In other words, pattern (a) will move towards the pattern that matches (a) best after blurred. We define this as "Blurred Matching" or BM.

In computer simulation, we may convolve a pattern with a Gaussian filter to get its blurred image For a practical algorithm of blurred matching, let function $f_a(x,y)$, fb(x,y) and $f_c(x,y)$ represent respectively pattern (a), (b) and (c) with identical mean brightness, and

$$G(x,y) = exp\left| -\frac{x^2+y^2}{2no^2} \right|$$

be the Gaussian filter, then the blurred patterns become

$$g_b(x,y) = f_b \otimes G$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_b(u,v)G(x-u, y-u)\,du\,dv$$

$$g_c(x,y) = f_c \otimes G$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_c(u,v)G(x-u, y-u)\,du\,dv$$

If the gravity centers of pattern (a), (b) and (c) locate at (x_a,y_a), (x_b,y_b) and (x_c,y_c), then we may examine how close (a) is to (b) or (c) by comparing the two distances:

$$D_{ab} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |f_a(x-x_a, y-y_a)$$

$$-g_b(x-x_b, y-y_b)|\,dx\,dy$$

$$D_{ac} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |f_a(x-x_a, y-y_a)$$

$$-g_c(x-x_c, y-y_c)|\,dx\,dy$$

The less the integral, the closer the two patterns. If D_{ah} is much greater than E, then we can be sure the apparent moving direction to most people would be (a) to (c), and vice versa. Experiments have showed this algorithm to be effective in predicting apparent motion in early stage vision. What if the two distances are close or equal? Perhaps we will have equally divided reports from subjects of (a) to (b) and (a) to (c). We may further assume two motion probability curves like shown in Fig. 3 Of course more work is needed to substantialize these points.

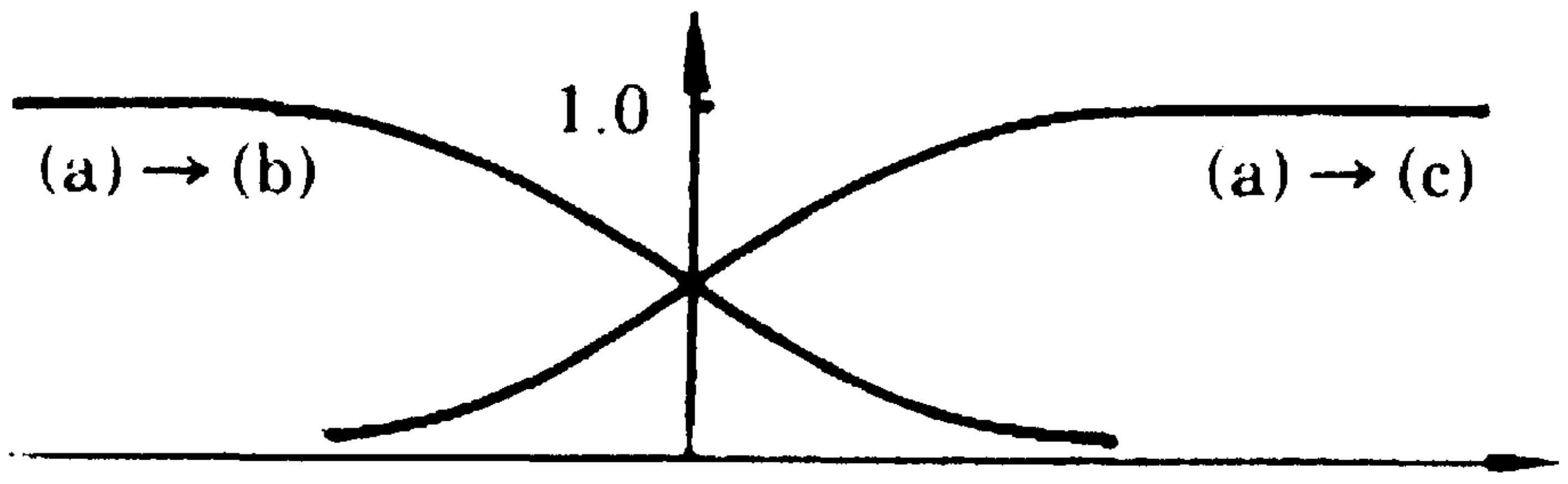

Fig. 3 Moving probabilities of (a) to (b) and (c) verses D_{ab}-D_{ac}

6 BM and Zero Crossing

It is notable the BM theory bears some resemblance with Marr's zero-crossing theory, where the same Gaussian filter *G(x, y)* is applied before the Laplacian operator V* to detect edges of scale o. This may lead to a misunderstanding that the former is a restatement of the latter. It is, however, not the case. In fact, apparent motion is not "computable", nor can it be explained in Marr's zero crossing for reasons below.

1). The V^2G operator in Marr's zero-crossing theory, applied to detect changes of intensity in a image, was not designed for the problem of apparent motion. Actually, the subject is almost not touched in his book "Vision".

2). There is a implied condition in Marr's intensity variation detecting theory, ie, human eyes have enough time to "scan" over the entire area of interest. This is essential for not discriminating against any part inside the area. The condition no longer hold true in perception of apparent motion, when human eyes was unable to move around in hundredsth of second, the time an image appears and disappears.

3). When applying the V-G operator, what we detect from an image is edges of various scales. As stated by Ramachandran, edges play little role in apparent motion. This indicates that such approach is appropriate for dealing with apparent motion On the other hand, edge detecting by V^2G operator is not feasible in the circumstances. We will go into some details on the point.

Marr pointed in his "Vision" that intensity variation occurs in a large range of scales. It can not be be all detected by a single operator. In applying the V^2G operator to an image, we are actually choosing m different parameters {Oj|i = I,..m} to form m Gassian filters, each is applied to a copy of the image to detect edges of a specific scale corresponding to o, The filtered results are then summed up. The process can be rewritten as

$$E = E_{o_1} + E_{o_2} + \ldots + E_{o_m},$$

$$E_{o_i} = \{f_i(x,y) | f_i(x,y) = \nabla^2 G(o_i) \otimes I\}$$

where

I: image,
E: set of all edges,
E_{o_i}: set of all edges whose scale corresponds to o_i in the ∇^2G operator.

If we assume that apparent motion follows the role of edge matching, we will have to answer the question that how these computation can be completed in human vision within miliseconds. The complexity of the computation and restricted processing time rule out any possibility of such edge matching.

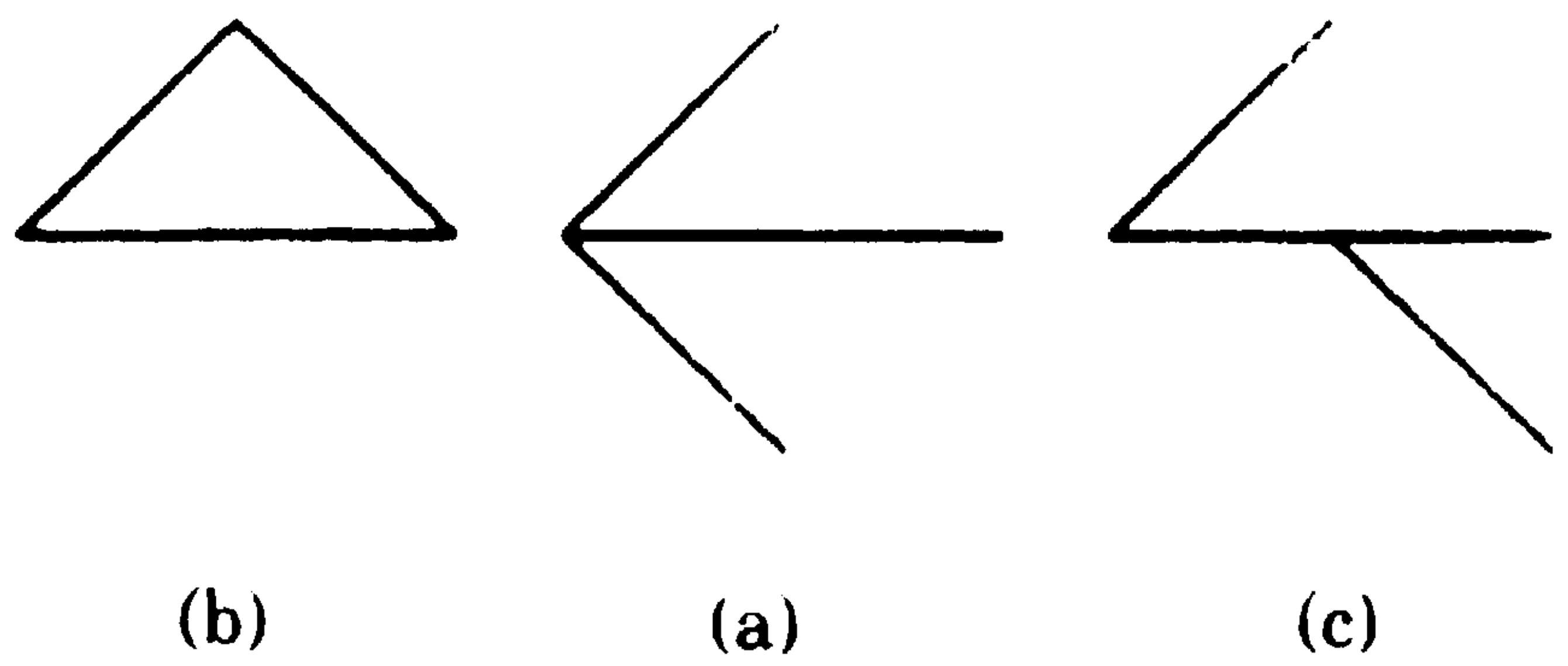
(b) (a) (c)

Fig. 4 Pattern (a) moves towards (c), which can not be explained as edge matching.

There is another difficulty that arises with edge matching Most people would agree that pattern (a) in Fig. 4 looks more like (b) than (c) on edges. Nevertheless, Chen's experiment showed that (a) would move towards (c).

Thus we conclude that although V*G operator seems similar to BM operator in their forms, they function differently in case of apparent motion, the the former is not applicable here.

References

[Chen, L., 1982a] Investigative ophthamology and visual science, vol. 22, 1982, p. 124

[Chen, L , 1982b] Topological structure in visual perception, Science, vol. 218, no. 12, 1982, p. 699

[Chen, L., 1982c] Bulletin of the psychonomic society, Sept., 1982

(Chen, L., 1983a| What are the units of figure perceptual representation, Studies in Cognitive Science, no. 22, (Sponsored by School of Social Sciences, University of California at Irvine), 1983

[Chen, L., 1983b] Topological invariants and apparent motion, Proceedings of the Fifth Annual Conference of the Cognitive Science Society, Rochester, 1983

[Marr, D., 1982] Vision, San Francisco, Freeman, 1982

[Ramachandran, 1986] VS., and Anstis, S.M., The perception of apparent motion, Scientific American, 1986

[Ullman, S., 1979] The interpretation of visual motion, Cambridge, Mass., The MIT Press, 1979