Perception, 1985, volume 14, pages 135-143

Perceptual organization in multistable apparent motion

Vilayanur S Ramachandran
Deparment of Psychology, University of California at San Diego, La Jolla, CA 92093, USA
Stuart M Anstis
Department of Psychology, York University, 4700 Keele Street, Downsview, Ontario M3J 1P3K, Canada
Received 17 October 1984

Abstract. Is motion perception based on a local piecemeal analysis of the image or do 'global' effects also play an important role? Use was made of bistable apparent-motion displays in trying to answer this question. Two spots were flashed simultaneously on diagonally opposite corners of a 1 deg wide square and then switched off and replaced by two spots appearing on the other two corners. One can either see vertical or horizontal oscillation and the display is bistable just as a Necker cube is. If several such bistable figures are randomly scattered on the screen and presented simultaneously, then one usually sees the same motion axis in all of them, suggesting the presence of field-like effects for resolving ambiguity in apparent motion.

While viewing a single figure observers experience hysteresis: they tend to adhere to one motion axis or the other and can switch the axis only by looking away and looking back after 10-30 s have elapsed. The figure can be switched off and made to reappear at some other random location on the screen and it is then always found to retain its motion axis. Several such demonstrations are presented to show that spatial induction effects in metastable motion displays may provide a particularly valuable probe for studying 'laws' of perceptual organization.

1 Introduction

Metastability is one of the most striking yet enigmatic aspects of perception (Rubin 1921; Gregory 1970; Attneave 1971; Julesz 1971; Rock 1975; Burt and Sperling 1981). Necker cubes (figure 1) and other bistable figures are often used to illustrate the point that perception is really an *opinion* on the state of affairs in the world rather than a passive response to sensory stimuli. One's perception of the cube changes dramatically as the mind hesitates between alternative three-dimensional (3-D) representations. Indeed, when viewing such bistable figures it is often hard to believe that something has not changed physically in the stimulus.

In this communication we report preliminary results of a series of experiments based on metastable apparent-motion displays such as that shown in figure 2b. If two spatially separated spots of light (figure 2a) are presented to the retina in rapid succession, the spot will appear to move from the first point to the second, as commonly seen in neon advertisement signs and traffic lights (Korte 1915; Anstis 1970; Julesz 1971; Kolers 1972; Ramachandran 1981). Figure 2b shows a matrix of four dots forming the corners

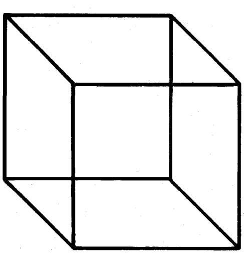


Figure 1. An outline drawing of a transparent cube. The mind hesitates between the two alternative 3-D representations.

of a square. The number by each dot refers to the time at which it is presented. Dots on two diagonally opposite corners are flashed first and then switched off, followed by two spots appearing simultaneously on the remaining two corners, with the procedure repeated in a continuous cycle. The two possible percepts, which are equally probable and mutually exclusive, are indicated in the diagram as percept 1 (vertical oscillation) and percept 2 (horizontal oscillation). Observers always report seeing one of these two percepts (the reader can confirm this by using the enclosed Apple disk; run the program labelled SQ1). A third theoretically possible percept is continuous clockwise (or anticlockwise) motion but this is almost never seen except under special conditions (see below).

This display is essentially bistable just as a Necker cube is. However, we found that it was difficult—almost impossible—to switch voluntarily from one percept to the other. Subjects almost always tend to persist in seeing one of the percepts (either vertical or horizontal oscillation) and can switch the axis of motion only by looking away and looking back after some time has elapsed. Exactly *how* much time has to elapse is one of the questions that we have raised in this report. Notice that what one adheres to is not one particular direction of motion (since dots on opposite sides of the square are actually moving in opposite directions), but to one particular axis—either vertical or horizontal. This tendency for the motion axis to persist indefinitely is a novel form of visual 'memory' that deserves further study. Also, metastable apparent-motion displays seemed to us to be less influenced by voluntary effort and attention than classical bistable displays and this might make them more suitable for laboratory investigation.

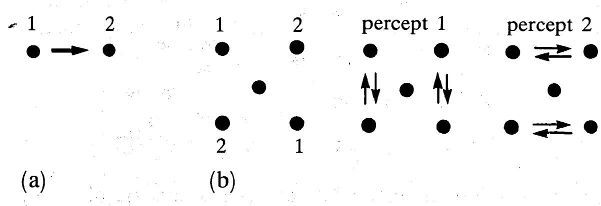


Figure 2. (a) Apparent motion seen between two light spots. Numerals (1 and 2) refer to the order of presentation. (b) The basic bistable display used in our experiment. The two possible percepts are shown (Apple disk program SQ1). Clockwise (or anticlockwise) rotation is almost never seen at short SOAs.

2 Method

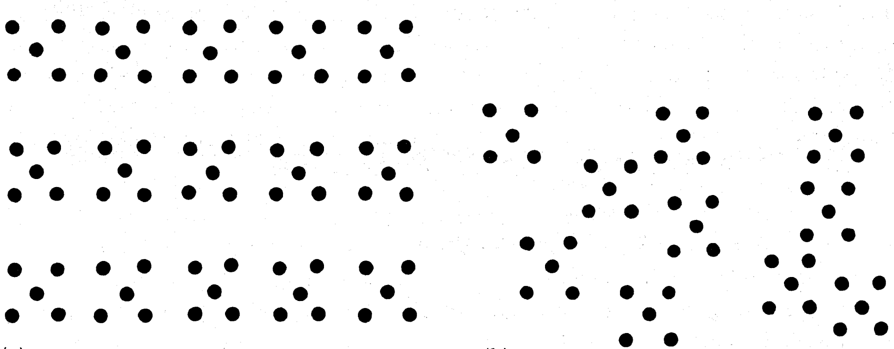
Our basic stimulus was the 'conceptual square' shown in figure 2b. This display (as well as subsequent ones described in this paper) was generated on a P-4 phosphor CRT with an Apple 2 microcomputer and viewed from a distance of 1 m. The sides of the square subtended 1 deg and the dots themselves were 3 min of arc in diameter. The numerals indicate the order of presentation. Note that only two dots (on diagonally opposite corners) are visible at any given instant. A hand-held potentiometer could be used to vary stimulus onset asynchrony (SOA) continuously over a wide range (25 to 500 ms). We varied SOA rather than the interstimulus interval (ISI) since the former is known to influence apparent motion more critically than the latter (Kolers 1972).

3 Experiments

3.1 *Experiment* 1

Our first experiment was concerned with the question of whether axis dominance is specific to a given region of the visual field or whether it spreads throughout the field. To answer this we generated several 'conceptual squares' such as that shown in figure 2b simultaneously on the screen (figures 3a and 3b; Apple disk programs MULTIREV and JITREV). The side of each square subtended 1 deg and the distance between squares was

1.5 deg. The speed of alternation was 3–4 frames s^{-1}. Would the subject perceive the same motion axis for all these displays or can different axes be simultaneously perceived? Twelve naive subjects (undergraduate students) viewing these displays reported that all the oscillating dot pairs seemed to 'lock in'—ie the displays always had the same motion axis. If one of the displays 'flipped' its motion axis, then all the displays invariably flipped with it simultaneously. This effect suggests the presence of a global 'field' for resolving ambiguity in bistable displays, for if the different displays were being processed independently there is no a priori reason why their oscillations should become synchronized. The presence of such a field is interesting for two reasons. First, it is known that if several Necker cubes or other similar reversible figures are simultaneously present in the visual field then they tend to undergo reversals more or less independently of each other (Long and Toppino 1981; Gillam 1981). This suggests that a field probably exists only for *apparent*-motion displays of the kind described here. Second, the effect seems to be somewhat inconsistent with the independence assumption of Ullman (1979) according to which apparent motion should be based on purely *local* computations and there ought to be no global field-like effects.



(a) (b)

Figure 3. (a) Multiple bistable displays presented simultaneously. Identical oscillations are observed in *all* the displays (Apple disk program MULTIREV; Paddle Controls SOA). As the SOA is made very slow, some degree of uncoupling can be obtained between displays that are widely separated. (b) The squares are randomly positioned and the same effect is seen which demonstrates that the presence of regular arrays is not necessary (Apple disk program JITREV; Paddle Controls SOA).

3.1.1 *The role of eye movements.* We considered the possibility that this tendency to synchronize the motion axes for multiple displays might arise at least in part from tracking eye movements. To rule this out we created two clusters of figures (such as that shown in figure 3a) side-by-side so that one of them was rotated in the frontoparallel plane by an angle of exactly 45° in relation to the other. The two sets of figures were separated by about 2 deg and were viewed simultaneously with one set presented to each hemifield. Three subjects viewed these and reported that the figures within each hemifield always had identical axes of motion, so that the axes for the two hemifields were separated by an angle of 45°. The subjects could also switch the motion axis for one cluster of figures (eg by temporarily occluding certain portions of it) without switching the axis in the other hemifield. Since one cannot track simultaneously in two directions, this result strongly suggests that eye movements cannot account for the effect reported here. For each cluster of figures a separate field must exist that synchronizes the elements within it.

3.1.2 *Dichoptic viewing.* To determine the site at which these global effects occur we used crossed polaroids to present the upper two rows of figure 3a to one eye alone and the third row to the other eye. Subjects fixated a cross (×) presented binocularly (unpolarized) between the second and third rows of the display and they reported that all the figures always had the same motion axis as in previous experiments. This result suggests that our global field-like effect must occur either at or beyond the site of binocular fusion in the brain (ie beyond layer 4 of the striate cortex).

3.2 *Experiment* 2

As pointed out earlier, observers tend to get 'stuck' to one or the other axis while viewing these displays and they usually have to look away for a few seconds in order to change the percept. To measure the duration of this visual persistence or 'set' we used six naive subjects who were unaware of the purpose of the experiment. They were initially familiarized with the two possible percepts by showing them the drawings labelled percept 1 and percept 2 (in figure 2b). Next, they were shown the original 'conceptual square' (figure 2b) and asked to report the axis of oscillation (vertical or horizontal). After 30 s of viewing, the display was switched off automatically, reappearing on the screen after various random intervals. The subject continued to fixate a cross on the screen and was instructed that after this period the axis of the display might or might not have changed. The period after which reports of direction were no longer significantly influenced by the initial perception was taken as a measure of visual persistence. Using this procedure we found that axis dominance persisted for at least 10 – 30 s in most subjects and sometimes even for several minutes. In fact, some observers tended to see the same axis for indefinitely long periods that were impractical to measure.

3.3 *Experiment* 3

We noted above (experiment 1) that for most observers the tendency to see a particular motion axis persists for *at least* 10 – 30 s after the display is switched off. We decided to explore the specificity of this 'memory'. If the initial preference for seeing a particular axis (say vertical) is set up in one region of the visual field, would this preference subsequently transfer to a distant location? To answer this question we created a display identical to that shown in figure 2b that was made to perform a 'random walk' on the screen. After every 3 – 4 cycles of oscillation it was switched off and made to reappear at some random location on the screen (figure 4; Apple disk program SQWALK1) after about 500 ms. Would the axis of motion within the square always remain the same even when the square moved to a new location? Six naive subjects viewed this display while fixating a central cross. They reported that the motion axis always remained the same even when the square moved to a new location, which suggests that the memory for the motion axis is not specific to retinal locus, ie it is not confined to the original locus where the axial preference was set up. This was true even when the jumps displaced the square across the midline.

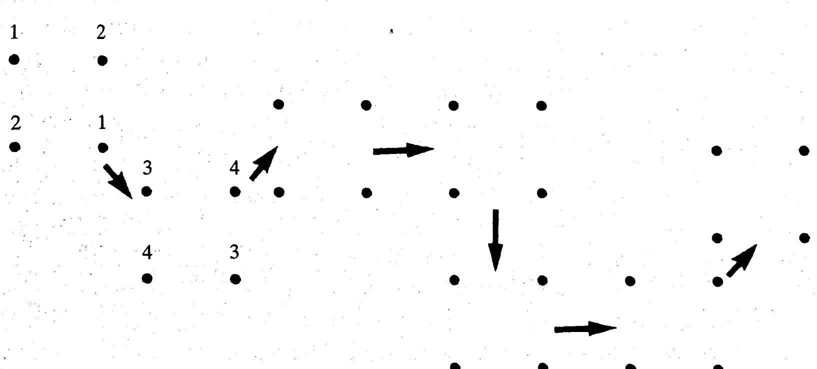


Figure 4. A bistable square performing a random walk on the screen (Apple disk program SQWALK1).

What would happen if a subject were to rotate his head by 90° in between two presentations? We have seen that if the square matrix (figure 2b) is switched off and switched on again in the same place after as long as 60 s, the subject tends to see the same axis. But, if he tilts his head between presentations, would the axis that is perceived now correspond to the original *retinal* axis or would it correspond to the original *phenomenal* axis? We found that the four subjects doing this experiment often reported seeing a percept that corresponded to the original retinal axis (80% of forty trials; ten trials per subject). This means that if the subject was originally seeing vertical oscillation, then on shutting his eyes and tilting his head by 90° and reopening his eyes he often saw vertical oscillation in relation to his head (ie horizontal in relation to the world). This is a surprising result, for it suggests that the 'memory' for the motion axis is only loosely coupled to constancy mechanisms that are involved in correcting for head orientation.

3.4 *Experiment* 4

We then explored the influence of an *unambiguously* moving array of dots in the surround on a single ambiguous square (similar to that shown in figure 2b) in the middle (figure 5). To our surprise we found that such unambiguous motion in the surround had only a slight influence on the perception of the central ambiguous square. For instance, one could easily 'switch' to a vertical motion axis for the ambiguous central display while all the dots in the surround were making purely horizontal excursions. It is interesting to contrast this result with the observation (experiment 2) that metastable motion displays always become synchronized in their direction of motion throughout the visual field.

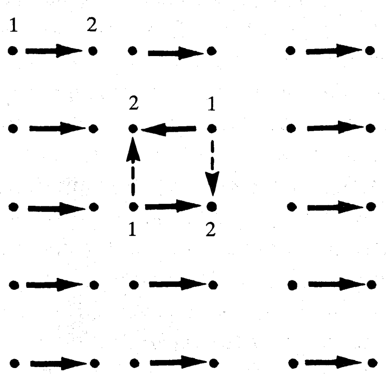


Figure 5. Demonstration that unambiguous motion in the surround with single pairs of dots has only a small effect on the perception of a bistable display in the center.

3.5 *Experiment* 5

In the previous experiments we have loosely referred to the motion axis of a 'conceptual square'. Obviously one needs a clearer notion of what *defines* the 'axis' of a figure. To answer this we began with the original display (figure 2b) again. After 3 cycles of oscillation the 'square' was switched off and replaced by a slightly rotated one (figure 6; Apple disk program TURN4). After another 3 cycles the square was again further rotated and this process of cogwheel-like rotation was continued in 20°–30° steps until a full 90° change in orientation had been achieved (subjects fixated a cross in the middle of the square). Surprisingly, we found that so long as the steps were small the motion axis always seemed to 'follow' the rotation of the square, so that if one began with vertical oscillation then one ended up with horizontal oscillation. This suggests that the motion axis is tied to whatever the visual system regards as the axis of the figure and will change its orientation depending on perceived changes in figural orientation.

To test this idea more directly we used the display shown in figure 7; this was produced by 'embedding' the original square (figure 2b) in a rectangular matrix of squares. The entire column of oscillating spots, A, was switched off after a few cycles of oscillation and then replaced by a horizontal row, B. The SOA for the oscillating dot pairs was 200 ms and the SOA between row A and row B was 300 ms. Subjects reported seeing the vertical row A clearly rotating to become the horizontal row B. The question is: what would be the direction of oscillation of spots within the horizontal row if one began with (say) vertical oscillation in the vertical row? Would it now be horizontal, corresponding to the figural axis, or would it continue to remain vertical, corresponding to the visual field axis?

Unfortunately the results of this experiment were not as clear-cut as those reported under experiments 1–4. There was, in general, a strong tendency for the motion axis to maintain a constant relationship with the *figural* axis rather than visual-field axis. Hence if one saw vertical oscillation for the vertical row, A, then when the row became horizontal, B, one would usually see horizontal oscillation of the dots; ie the motion of the dots would continue to remain parallel to the axis of the row as a whole. However, while this was *generally* true, occasionally subjects would report that the dots continued to oscillate vertically even when the whole row had now become horizontal; ie the oscillation of dots seemed to adhere to the visual-field axis rather than the figural axis. This suggests that while there is a strong coupling between motion axis and figural axis there is also a competing tendency to preserve the original visual-field axis. The reason for this competing tendency is obscure but it is obviously related to the global field effect reported under experiment 1.

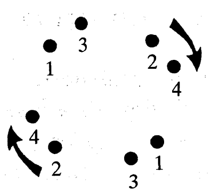


Figure 6. Cogwheel-like rotation of the bistable square. The square retains its motion axis (Apple disk program TURN4).

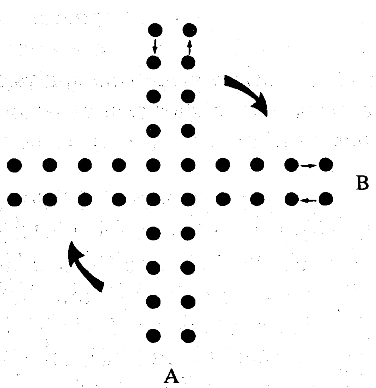


Figure 7. When a vertical row of oscillating dots, A, is replaced by a horizontal row, B, the motion axis tends to adhere to the figural axis (see text).

3.6 *Experiment* 6

We noted earlier that although subjects generally report oscillation of dots, one should *in principle* be able to see either continuous clockwise or anticlockwise motion of the dots.

At the SOAs used in our earlier experiments (100–200 ms) this was impossible to achieve even through intense mental effort. However, we made the curious observation that when the SOA was made sufficiently long (> 350 ms) one could 'will' oneself to occasionally see clockwise or anticlockwise motion. We asked three subjects to do this, and while they watched the display we gradually lowered SOA. When the SOA was less than 300 ms, the subjects reported that they could no longer 'hold' this percept and that they had once again reverted to seeing oscillation (either vertical or horizontal). This oscillation could not be switched back to continuous rotation unless the SOA was once again lengthened. We conclude from this experiment that axis dominance remains absolute and unambiguous at SOAs smaller than 250 ms and that it becomes 'coupled' with the effects of 'will' and attention only at long SOAs. Hence using short SOAs may be a good way to study metastability without worrying about contamination from 'higher centres'. Alternatively, we could actually study the influence of those centers by using our displays. A glass of champagne, for instance, might significantly increase the SOA at which one started seeing oscillation instead of rotation.

4 Discussion and conclusions

Experiments with our basic four-dot metastable apparent-motion display have led us to the following observations:

(i) Dots in diagonally opposite corners of an imaginary square are flashed first and then switched off and replaced by dots appearing on the remaining two corners. If this procedure is repeated in a continuous cycle, observers always report seeing either vertical or horizontal oscillation, and they almost never see the equally plausible percept of continuous clockwise or anticlockwise rotation (except at very long SOAs, > 300 ms).

(ii) At short SOAs the direction of motion-axis dominance cannot be influenced by voluntary effort. As the SOA is increased there is an abrupt transition to a state where one can voluntarily change the direction of axis dominance. It may be that when events are too rapid in the module that computes apparent motion, it cannot be coupled with the 'will' mechanism (which may have a long time-constant).

(iii) This axis dominance effect spreads throughout the visual field; ie if one views several displays simultaneously, then one always sees the *same* motion axis for all the displays, although there is no a priori reason why this should happen. An apparently global effect of this kind could be based either on cross talk (ie lateral interactions) between independent motion-detecting modules which perform a piecemeal analysis of the visual field, or alternatively there might be feedback from higher centers which introduces the same directional bias in all the modules. Further experiments are now in progress to distinguish between these two possibilities.

(iv) If the display is switched off there is a 'memory' for axis dominance that lasts for at least 10–30 s in most observers and for a much longer duration in some. This memory transfers quite readily from one location in the visual field to another distant location, which suggests that it is not specific to the particular retinal area that was 'trained'. It is interesting to contrast this with the observation (Ramachandran 1976) that certain other kinds of memory, such as memory for global stereopsis, do *not* transfer from one visual hemifield to the other.

(v) Axis dominance is also tied to 'figural axis'. If the square matrix is rotated in steps like a cogwheel then the motion axis seems to follow this rotation faithfully. Hence it may turn out that motion-axis dominance is a convenient probe for finding out what rules the visual system uses for establishing *figural* axis in a given situation.

During the earlier part of this century Gestalt psychologists (Köhler 1929; Koffka 1935) used both static and dynamic displays to try and discover the 'laws' of perception and they believed that the interpolation process in apparent motion was particularly revealing of fundamental organizing principles. Unfortunately, although they had many valuable insights to offer, they had the bad habit of inventing descriptive phrases such as 'prägnanz' or the 'law of common fate' which were often offered as *explanations* rather than mere labels. In a different context the British biologist Medawar (1967) has referred to explanations of this kind as 'analgesics'; for they seem to "dull the ache of incomprehension without removing the cause...".

As an example consider our basic ambiguous display (figure 2b). One could argue that seeing continuous rotation (clockwise or anticlockwise) has more prägnanz than seeing oscillation, since a circle is a good gestalt (as often emphasized by the Greeks). However, as we noted earlier, the preferred percept is one of oscillation. Of course, one could always argue that oscillation *also* has prägnanz but this makes the whole doctrine utterly untestable.

The displays we have generated and the effects we have explored are certainly very similar to the perceptual grouping effects of Gestalt psychology (eg figure 8) but there are several important differences. First, the effects reported here are much more compelling and also seem relatively immune from subjective factors such as 'will' and 'attention' at least at short SOAs and this might make them more easily quantifiable. A second difference is that changing the display in certain ways seems to produce systematic measurable changes in what is perceived (eg see Ramachandran and Anstis 1983). Hence metastable motion displays may give us a better hold on those elusive laws of perceptual organization that were so keenly sought after by Gestalt psychologists.

Perception may be regarded as essentially a process taking advantage of certain invariances in order to 'home in' on a unique solution. The basic strategy for studying perception therefore ought to be to confront the visual system with ambiguous displays and to ask: what *rules* does the system use to resolve ambiguities in a given situation? The purpose of this preliminary report has been to demonstrate that ambiguous motion displays may be particularly suitable for discovering these rules.

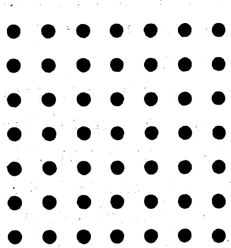


Figure 8. An array of dots which can be mentally organized into either vertical or horizontal rows.

Acknowledgements. We thank Ebbe Ebbesen and Ken Wexler for providing facilities and John Yellott, Francis Crick, and Bela Julesz for helpful comments.

References

Anstis S M, 1970 "Phi movement as a brightness subtraction process" *Vision Research* **10** 1411 – 1430

Attneave F, 1971 "Multistability of perception" *Scientific American* **225** 62 – 71

Burt P, Sperling G, 1981 "Time, distance and feature trade-offs in visual apparent-movement" *Psychological Review* **88** 171 – 195

Gillam B, 1981 "Separation relative to length determines the organisation of two lines into a unit" *Journal of Experimental Psychology: Human Perception and Performance* **7** 884 – 889

Gregory R L, 1970 *The Intelligent Eye* (London: Weidenfield & Nicolson)

Julesz B, 1971 *Foundations of Cyclopean Perception* (Chicago, IL: University of Chicago Press)

Koffka K, 1935 *Principles of Gestalt Psychology* (New York: Harcourt)

Köhler W, 1929 *Gestalt Psychology* (New York: Liverright)

Kolers P A, 1972 *Aspects of Motion Perception* (New York: Academic Press)
Korte A, 1915 "Kinematoskopische Untersuchungen" *Zeitschrift für Psychologie* **72** 193 - 296
Long G M, Toppino T C, 1981 "Multiple representations of the same reversible figure" *Perception* **10** 331 - 334
Medawar P B, 1967 *The Art of the Soluble* (London: Methuen)
Ramachandran V S, 1976 "Learning-like phenomena in stereopsis" *Nature (London)* **262** 382 - 384
Ramachandran V S, 1981 "Perception of apparent movement" in *Studies in the Cognitive Sciences* volume II, Bulletin issued by the School of Social Sciences, University of California, Irvine, CA 92717
Ramachandran V S, Anstis S M, 1983 "Extrapolation of motion path in human visual perception" *Vision Research* **23** 83 - 86
Rock I, 1975 *An Introduction to Perception* (New York: Macmillan)
Rubin E, 1921 *Visuell wahrgenommene Figuren* (Copenhagen: Gyldenalske Boghandel). Reprinted as "Figure and ground" in *Readings in Perception* eds D C Beardslee, M Wertheimer (1958, Princeton, NJ: Van Nostrand) pp 194 - 203
Ullman S, 1979 *The Interpretation of Visual Motion* (Cambridge, MA: MIT Press)