

MOTION CAPTURE ANISOTROPY

V. S. RAMACHANDRAN

Department of Psychology, University of California at San Diego, La Jolla, CA 92093, U.S.A.

and

PATRICK CAVANAGH

Departement de Psychologie, Universite de Montreal, Montreal, Quebec, Canada H3C 3J7

(Received 21 August 1985; in revised form 18 April 1986)

Abstract—Two uncorrelated random dot patterns were superimposed and alternated to produce dynamic incoherent noise. When a low spatial frequency sine wave grating was optically superimposed on this noise and moved in step with the alternation of the two frames, the incoherent motion was masked and all the dots were seen to adhere to the grating and to move with it as a single rigid sheet ("Motion Capture"). Over a wide range of displacements subjects could not discriminate uncorrelated noise which was "captured" from correlated noise patterns which moved physically in the same direction as the grating. In fact the motion signal from the low frequency grating was even strong enough to overcome signals from two *correlated* random dot patterns which moved in the opposite direction. Capture was not as strong if the direction of dot motion was orthogonal to the direction of grating motion. We conclude that in any dynamic visual scene the motion of certain salient features in the image tends to dominate our perceptual experience. The signal from low frequencies masks or inhibits the signal from the high frequencies. Since the latter now have no motion signal of their own they are assumed to move with the low frequencies. Thus, motion capture suggests an important biological role for long-range apparent motion: the process serves to preserve continuity of object identity while at the same time eliminating spurious motion signals that arise from finer image features. In this manner the visual system solves the "correspondence problem" without benefit from either computation or cognition.

Motion Random-dots Correspondence problem Apparent motion

When an object is suddenly displaced over a large distance, the visual system is presented with an interesting problem. The overall outline of the object may produce a strong motion signal but the inner detail of the object may produce conflicting signals or even no signal at all. As Braddick (1974) has shown, there is an upper displacement limit (D_{\max}) over which the coherent motion of a textured surface can be perceived. If the inner texture of the object is displaced a distance greater than this limit, ambiguous signal will be produced (Chang and Julesz, 1983; Ramachandran and Anstis, 1983). If the object jumps so that there is no overlap whatsoever between its first and second positions, the texture alone can produce no signal at all as no process exists in the brain to verify the correspondence between the initial texture and the final texture across a large, nonoverlapping displacement. And yet, in such a case we do not see the texture simply disappear from the starting position and then reappear at the final position. The texture appears to jump with the object even though it is unlikely that any visual process is capable of asserting that it did so.

If the object were to jump over a smaller displacement, the texture could produce signals of local motion because of random pairings of neighboring points in the initial and final distributions. These motion signals are in conflict with the motion signal from the object's border however. How is this conflict resolved? Will a leopard's spots appear to move about randomly if the leopard takes a jump forward short enough so that many of the leopard's spots overlap in the initial and final positions but far enough so that it is beyond the displacement limit for correct interpretation of the dots' motion? Our evidence says that they will not. The spots will appear to move along with the outline even though in the absence of the outlines the dots would produce incoherent motion. We have called this phenomenon motion capture (Ramachandran, 1981; Ramachandran and Anstis, 1983; Ramachandran and Inada, 1984; Ramachandran and Inada, 1985). In a stimulus with conflicting motion information, the motion signal of the more salient feature captures the motion of the other features; its motion is attributed to the whole object. When incoherent

dot motion is superimposed on a low frequency sinewave that is alternating in position, the dots appear to be glued onto the sinewave and move with it. The independent, local motions of the dots are not longer seen.

The purpose of this paper is to determine whether a salient feature can capture and override a coherent motion signal as well as an incoherent one, to determine if this capture occurs more readily along directions colinear with the motion of the salient feature or orthogonal to it and finally to examine the importance of spatial frequency content.

EXPERIMENT 1

Method

The display (see Fig. 1) subtended 8° by 8° on a CRT screen with a central dividing strip of $1^\circ 30'$ high separating the upper and lower test fields. A $10'$ wide fixation spot was located in the center of the dividing strip. A sinewave grating of 1 cycle per degree (c/deg) and 40% contrast was present in both upper and lower test fields. Random dot arrays were superimposed on the sinewaves with each dot subtending $3.75'$ by $3.75'$ and having a contrast of 40% with the luminance of the portion of the sinewave immediately surrounding it. There was approximately 600 dots in each of the two fields (7% density).

Each trial was presented as a continual alternation between two frames with 500 msec SOA. (The ISI was zero in this and in all subsequent experiments.) The dots above the fixation point were correlated in successive frames and displaced horizontally or vertically by varying amounts ($3.75'$, $7.5'$, $11.25'$ and $15'$ of arc for observers P.C. and E.C.; $1.875'$ $15'$ for V.S.R.). Motion of the dots on any given trial was either horizontal only. The grating on which the dots were superimposed alternated between two horizontal positions in successive frames and the distance over which the grating moved was under the observer's control. Below the fixation point the dots were uncorrelated in successive frames but the grating motion was identical.

The observer's task was to indicate when the top and bottom fields appeared dissimilar. They were instructed not to scrutinize and compare individual dots. Each trial began with the grating making a 120° phase jump. If the upper and lower fields appeared the same, the subject reduced the grating jump size until he or she could see a difference between the two. The grating jump at which the difference in the dot

motions of the upper and lower fields became visible was taken as the measure of the strength of motion capture. Four readings were taken for each value of dot displacement and orientation of dot motion. The order of the conditions was randomized and the location of the coherent and incoherent dot fields (top or bottom) was assigned randomly for each trial.

Results

Figure 2 (*left-hand column*) shows the results of this experiment. Note that for combinations of grating and dot jumps falling above and to the right of the line joining the data points, the subjects had difficulty in discriminating coherent from incoherent dot motion. In fact, in these cases the subjects perceived the dots as moving with the sinewave grating in both the upper and lower fields; in effect, the dot motion was "captured" by the grating motion. For combinations to the left and below the lines, motion capture did not occur. The subjects could see the individual dot motions and discriminate the upper from the lower field. Discrimination was still possible at the maximum dot jump, $15'$, indicating that this jump was below D_{\max} for these dot fields. Also, the data show that there was no systematic difference in the magnitude of capture whether the dot field and the sinewave grating moved *with* each other or *against* each other (but see Experiment 4, below).

We conclude that as soon as the grating displacement becomes great enough there is a tendency for the motion signals arising from the matches between individual dots to be ignored. The visual system acts as if it were attributing the grating motion to the dots themselves.

The data demonstrate a significant anisotropy in the strength of motion capture for all three subjects. A larger grating displacement was required to capture vertical dot motion than horizontal dot motion. In particular, the smallest vertical dot motion was never captured even at the largest grating excursions used in the tests (a 120° phase jump, larger jumps have not been found to improve capture, Ramachandran and Inada, 1985; and at 180° the motion becomes ambiguous).

Finally, for the conditions in which the dots in the coherent field were moving in the same direction as the grating and had the same displacement value, the dots were, in fact, glued to the grating. These special instances did not lead to any perceptual advantage, however. These cases were no more or less easy to "capture"

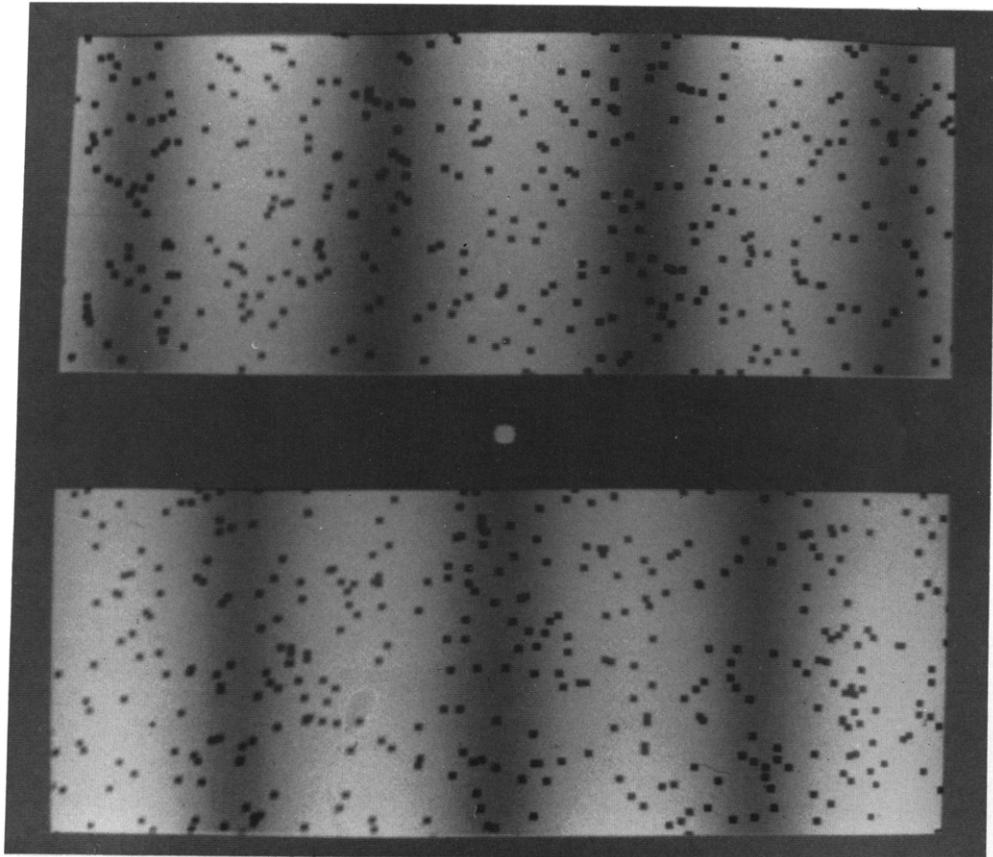
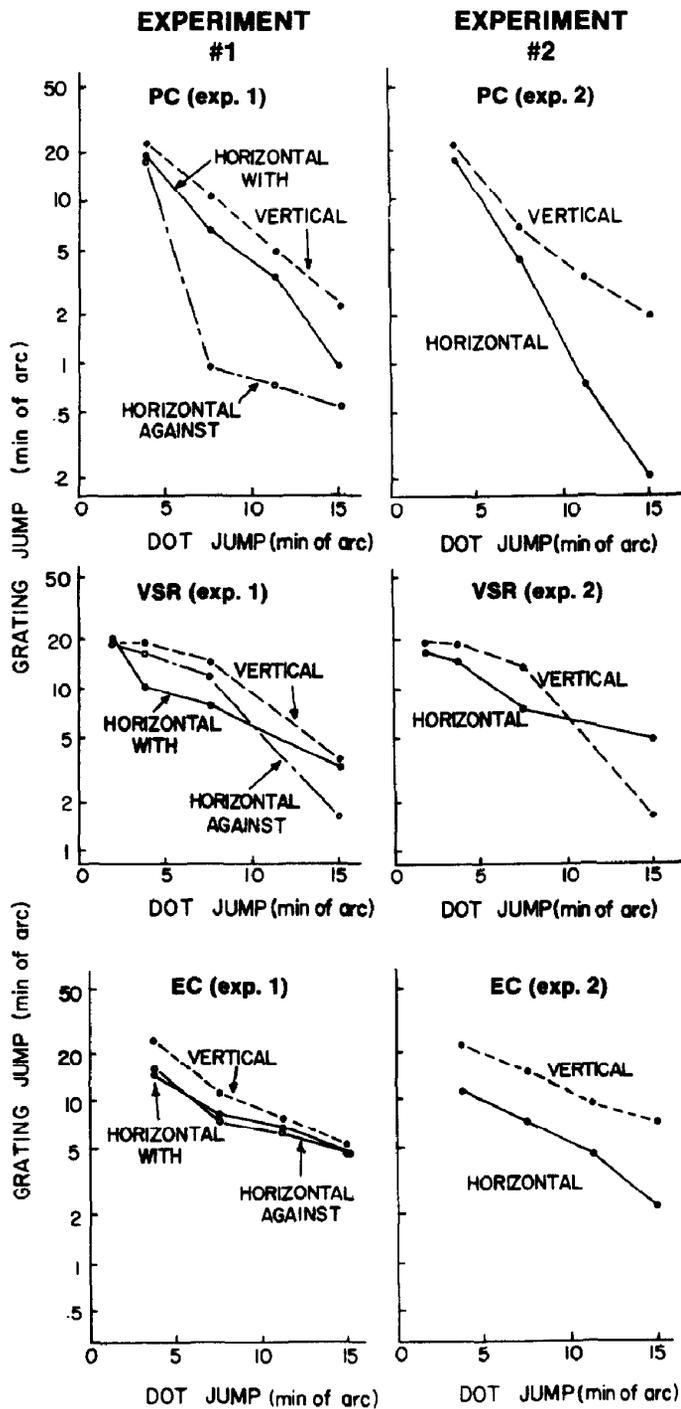


Fig. 1. An example of the stimuli used in the experiment. Subjects fixation on + on the central horizontal strip. Their task was to gradually reduce the excursion of the grating until the two fields looked dissimilar.



than cases where the dots and grating moved independently. In other words, for dot and grating displacements of 10' and 15', the field where the dots were physically attached to the grating could not be discriminated from the field where the dots were moving incoherently. The dots in both fields appeared equally stuck to the moving gratings.

In this experiment, subjects had to discriminate between coherent and incoherent dot motion. For grating motions beyond a certain value the two dot motions could not be discriminated because both were captured by the grating and looked similar. For shorter grating excursions, one or other of the dot motions would be released from capture and the two fields could be discriminated. Although it would appear likely that it was the coherent dot motion that was first released, it is not possible to determine if this was true from the results of this experiment alone.

EXPERIMENT 2

Method

The procedure for the second experiment was identical to that used for the first except that the dot motion in both fields was coherent, but in opposite directions. The grating motion was always horizontal; the dot motion could be vertical or horizontal. When it was horizontal, it was in the same direction as the grating motion in one field and opposite to it in the other. The direction of grating motion was identical for both fields. Subjects gradually reduced the grating displacement until they could just discriminate the two fields.

Results

The results (Fig. 2, *right-hand column*) were quite similar to those of Experiment 1. The shorter the dot jump, the farther the grating had

to jump to capture it, implying that shorter jumps produce stronger motion signals. Vertical dot motion was again more resistant to capture.

The similarity of these data to those of Experiment 1 indicate that in that experiment, it was the coherent motion that first escaped from motion capture allowing discrimination of the two test fields. This was to be expected from the data of Experiment 1 as incoherent motion could be considered motion beyond D_{\max} and it is evident in Fig. 2 (*left*) that the closer the coherent motion jump was to D_{\max} , the easier it was to capture. The first field to be released from capture would therefore be the coherent field.

EXPERIMENT 3

Next, we investigated the spatial frequency characteristics of the motion capture effect. Rather than superimposing dots on a sinewave grating, we superimposed a second sinewave grating of a different frequency and/or orientation. The two spatial frequencies had to differ by at least a factor of three for the stimulus motion to be unambiguous. With a frequency ratio of two, the various local brightness profiles produced by different relative offsets of the two gratings made the motion very ambiguous and confusing.

Method

The procedure was identical to that for Experiment 2 except that rather than a dot field, a second sinewave grating of 2.67 c/deg and 40% contrast was added linearly to the first sinewave, moving in the same direction in one field and the opposite direction in the other field. The first sinewave had 40% contrast as well and four different spatial frequencies were investigated: 0.27, 0.44, 0.59 and 0.88 c/deg. These fre-

Fig. 2. (*Facing page*) Depicts results of experiments 1 and 2. In Experiment 1 (*left-hand column*), dots in one field were uncorrelated in successive frames and in the other field they were correlated and displaced horizontally (solid line) or vertically (dotted line). A vertical sine-wave grating of 1 c/deg and 40% contrast was superimposed on both fields. Subject's task was to gradually reduce the displacement of the grating until the dots were released from capture so that the two fields became discriminable. Above and to the right of the graph the dots were captured and the two fields were indiscriminable. (Each datum point is the mean of four readings.) Note that vertical dot motion is more difficult to capture. Horizontal motion of the correlated dot field in the same direction as the grating (WITH) does not appear to be either easier or harder to capture than motion in the opposite direction (AGAINST). In Experiment 2 (*right-hand column*) the dots were always correlated in successive frames both above and below the central divider. The grating moved in the same direction in both fields but the dots moved in opposite directions. For instance, for vertical dot motion (evenly interrupted line) the dots in the two fields moved either simultaneously towards or away from the central dividing strip. As in Experiment 1, subjects reduced the grating displacement until they could just discriminate the two fields.

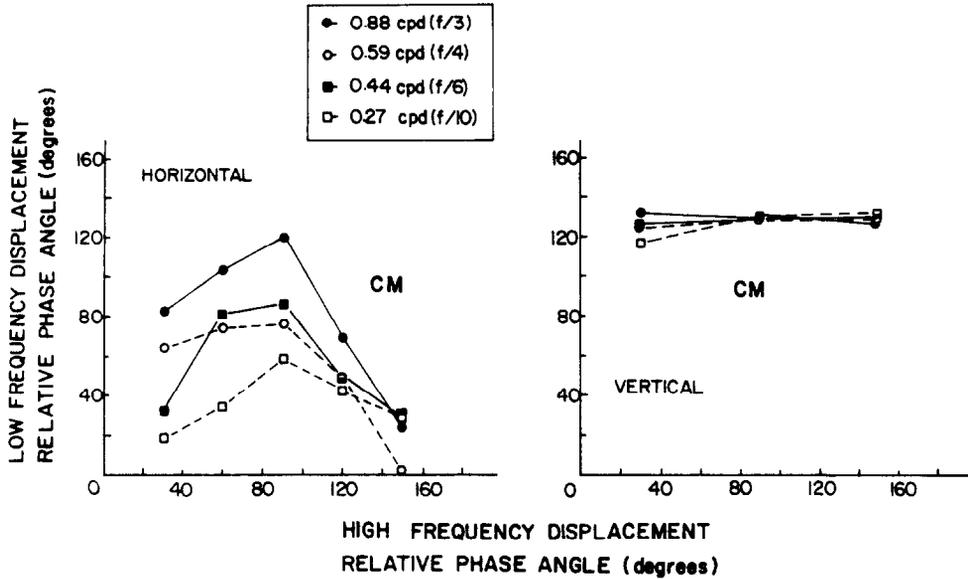


Fig. 3. (*Horizontal*) Depicts capture of a high spatial frequency grating by a low frequency grating. Procedure identical to Fig. 2 except that instead of a dot-field, a second sinewave grating (2.67 c/deg; 40% contrast) was added to the first sine-wave. The high frequency grating moved in the same direction as the low frequency grating in one field in the opposite direction in the other field. Readings were taken for each of 5 horizontal displacements of the high frequency grating. Capture was observed for high and low frequency displacement combinations *above* the plotted lines. Four different low-frequency gratings were used (0.27, 0.44, 0.59 and 0.88 c/deg). Note that the lowest frequency (0.27 c/deg) is most effective in producing capture. (*Vertical*) Similar to the previous graph except that the motion of the high frequency grating was *vertical*. (The grating moved in opposite directions in the two fields; either towards or away from the central dividing strip.) Motion of the low frequency grating was horizontal and identical in the two fields. The high frequency grating was necessarily oriented horizontally. Note that capture is considerably reduced compared to the previously graph; the vertical motion is visible even at the largest excursion of the horizontal grating.

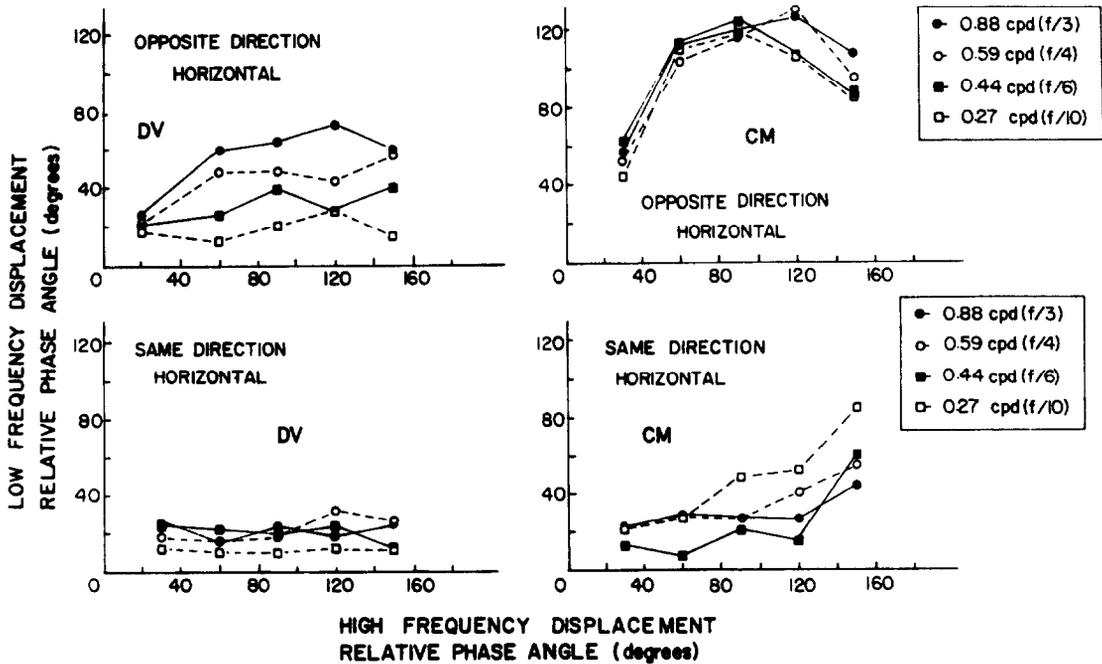


Fig. 4. In one field the high-frequency grating was stationary and superimposed on a moving low-frequency grating. In the other field the grating moved either *against* (opposite direction) or *with* the low-frequency grating (same direction). Note that a grating moving in the opposite direction was more difficult to capture. Data for two subjects are shown separately.

quencies had harmonic relations of 1/10, 1/6, 1/4 and 1/3 to the high frequency grating.

The observer's task was to indicate when the top and bottom fields appeared qualitatively different. Each trial began with the low frequency grating making a 120° phase jump, if the upper and lower fields appeared the same, the subject reduced the jump size of the low frequency grating until he could see a difference between the two. The grating jump at which the motion of the high frequency grating became visible was taken as the measure of the strength of motion capture. Four readings were taken for each of the five horizontal displacements (phase angles of 30, 60, 90, 120 and 150°) and three vertical displacements (phase angles of 30, 90 and 150°) of the high frequency grating. For the vertical displacements, the high frequency grating was necessarily oriented perpendicular to the low frequency grating.

Results

The results (Fig. 3, *left*) show that 90° displacements of the higher frequency grating were the hardest to capture (needed the largest low frequency grating excursions). The smallest (30°) and largest (150°) displacements tested were the easiest to capture. These data are consistent with those of Nakayama and Silverman (1985) showing that 90° phase jumps produce the largest motion signal.

The lower the low spatial frequency, the more effective it was at capturing the high frequency grating. This did not scale to the physical as opposed to phase angle jump size however.

The capture effect again showed an anisotropy for the grating stimuli that was even more marked than for the dot stimuli. Vertically moving gratings were almost never captured by horizontally moving gratings (Fig. 3, *right*).

EXPERIMENT 4

We were next interested in determining whether there was a difference in motion capture for gratings moving in the same vs opposite directions.

Method

The procedure was identical to that for Experiment 3 with two exceptions. First, the high spatial frequency grating in one field was stationary while in the other field it either moved with or against the low frequency grating. The fields were alternated randomly over trials. Sec-

ond, only horizontally moving gratings were tested.

Results

As Fig. 4 shows, gratings moving the same direction were strongly captured. Gratings moving in the opposite direction were weakly captured, showing approximately the same results as were seen in the previous experiment. Evidently the data of the previous experiment were determined largely by the release from capture of the grating moving in the direction opposite to the low frequency grating. The effect of spatial frequency is much less noticeable in the second subject (C.M.).

EXPERIMENT 5

Finally, we evaluated the ability of a higher spatial frequency to capture the motion of a lower spatial frequency grating.

Method

The procedure was similar to that for Experiment 3 except that the role of the high and low spatial frequency gratings were exchanged. The spatial frequencies used were the same as in Experiments 3 and 4.

The observer's task was to indicate when the top and bottom fields appeared qualitatively different. Each trial began with the high frequency grating making a 120° phase jump. If the upper and lower fields appeared the same, the subject reduced the jump size of the high frequency grating until he could see a difference between the two. The grating jump at which the motion of the low frequency grating became visible was taken as the measure of the strength of motion capture. Four readings were taken for each of the five horizontal displacements (visual angles of 1, 2, 3, 4 or 5') and three vertical displacements (visual angles of 1, 3 or 5') of the low frequency grating. For the vertical displacements, the low frequency grating was necessarily oriented horizontally, perpendicular to the high frequency grating.

Results

The physical displacements of the low frequency gratings were converted to phase angle of each grating in Fig. 5. Independently of the spatial frequency, movement became visible at a fixed relative phase difference between successive positions of the grating; approximately 5° for the horizontal motion and 3° for the vertical.

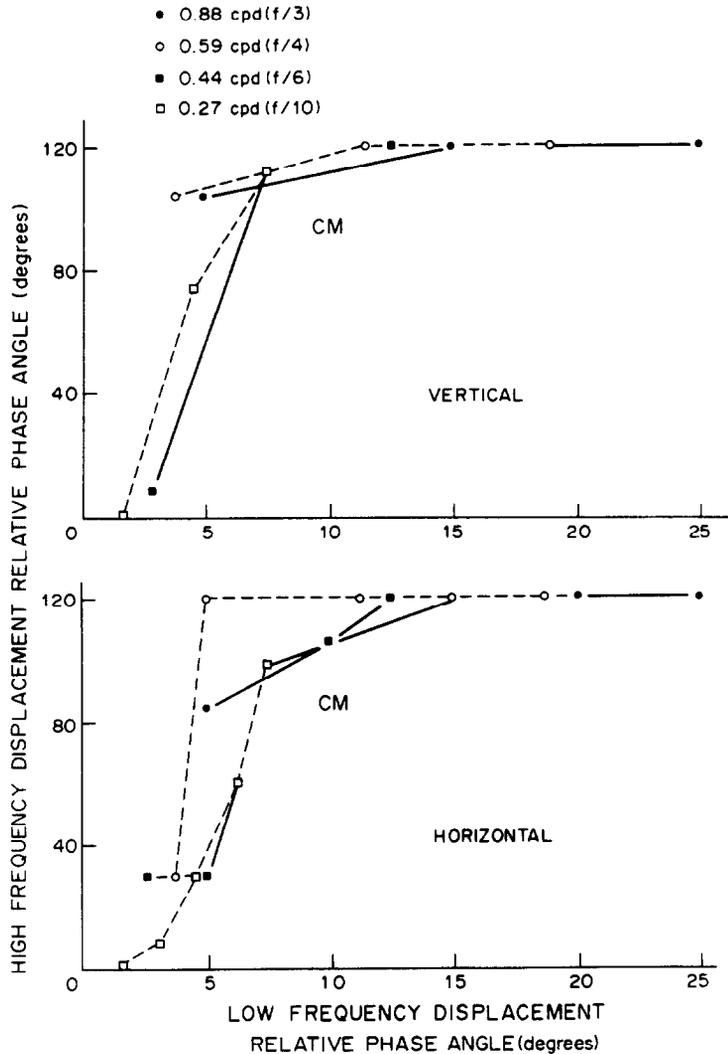


Fig. 5. Displacement of high spatial frequency grating (2.67 c/deg) necessary to capture motion of low spatial frequency grating. The opposite directions of motion of the low frequency grating in the upper and lower fields were not visible (i.e. were "captured") for displacement combinations above and to the left of the plotted lines.

Thus while there was some slight orientation anisotropy, there was little effect of spatial frequency.

On the other hand, the low frequency motion that the high frequency grating can capture is smaller by at least an order of magnitude than when the roles were reversed as in Experiments 3 and 4. The data reveal a major asymmetry in the strength of motion capture; low frequency gratings are much more capable of capturing high frequency gratings than vice versa.

CONCLUSIONS

The experiments reported in this paper establish "motion capture" as a robust perceptual

effect and allow us to draw three major conclusions:

(1) Over a wide range of conditions subjects could not discriminate incoherently moving spots that were "captured" by a moving grating from dots that moved coherently along with the grating. The same two patterns could be discriminated instantly if the grating motion was stopped suddenly. This suggests that as soon as the grating displacement becomes large enough, the motion signal derived from it inhibits the motion signals arising from the individual dots and the visual system simply attributes the grating motion to the dots themselves.

(2) Horizontal grating motion captured coherent apparent motion of correlated dots that

moved either vertically or horizontally. However a significantly larger grating displacement was required to override vertical dot motion than horizontal dot motion—suggesting that the capture effect is direction specific.

(3) Experiments with gratings confirmed this anisotropy and also suggested that low spatial frequencies were far more able to capture high frequencies than vice versa. However, a high frequency grating was easier to capture if it moved *with* rather than *against* the direction of the low frequency grating.

Although we had instructed our subjects to carefully fixate the central fixation spot on the dividing strip we were concerned that the motion capture illusion may at least in part have arisen from tracking eye movements. To rule out this possibility we set up a demonstration in which the grating performed a to and fro rocking action instead of moving horizontally. When this rotating grating was superimposed on dynamic noise the dots adhered to the grating and appeared to move with it. Observers fixating the center of the display reported that dots which were on different sides of the fixation spot moved simultaneously in opposite directions. Further, when we gradually reduced the grating excursion a point was reached when the dots were “released” from capture as in Experiment 1. These results suggest that tracking eye movements cannot be invoked to explain the motion capture illusion. (Also see Ramachandran and Inada, 1985).

Despite over a century of research on apparent motion there have been very few attempts to discover what its *function* might be, and in fact the phenomenon is often relegated to the role of an amusing classroom demonstration. Our results suggest an important biological role of “long-range” apparent motion; i.e. it serves to keep track of the identity of a moving object while at the same time eliminating spurious motion signals that arise from finer image features. Such a process may be especially valuable in three situations; (a) when a splotchy object (e.g. a leopard) makes a rapid excursion; (b) when such an object’s trajectory is momentarily occluded by an opaque object such as a tree-trunk (Ramachandran, 1981, 1984) and (c) when the texture on the objects surface changes as it moves (e.g. when shadows of leaves are falling on its surface).

Although the mechanisms underlying motion capture require further study, we tentatively postulate the following sequence of events to

account for the illusion: (1) Motion is first extracted separately from different spatial frequency bands. (2) For large excursions motion signals from low spatial frequencies (and other salient features) mask or *inhibit* the signals derived from finer image features—a process that serves to eliminate spurious or incoherent motion signals. (3) The inhibition of motion signals from the high frequencies does not cause them to appear stationary; they are in fact seen to jump as though they were glued to the low frequency grating. This suggests that when the high frequencies have no motion signal of their own they are assumed to move with the low frequencies, *by default*. The general rule is that if there are no motion signals from some frequency bands and strong signals from another (lower) frequency band the signal from the latter is spontaneously attributed to the former. Thus “masking” and “capture” should be regarded as complementary, rather than incompatible descriptions of the same phenomenon.

The disadvantage with the masking terminology is that it fails to emphasize the functional role of these phenomena and also has connotations of “critical band”. (In contrast to what we have observed in our displays masking usually occurs between *adjacent* frequency bands.) The major advantage with the scheme, however, is that it implies that an apparently “high-level” process such as preserving continuity of object identity during the motion may be achieved through a simple interplay of motion signals derived from different spatial frequency bands. This scheme is very different from both the iterative “minimum distance” algorithm that has been proposed for motion correspondence (Ullman, 1979) and from cooperative algorithms that have been proposed for stereopsis (Marr, 1982).

What is the functional significance of motion capture? We believe that the process might help solve the so-called “correspondence problem”. Consider successive views of a leopard jumping rapidly from branch to branch on the treetops while pursuing one of our arboreal ancestors. For long jumps of the leopard, the excursion of dots is beyond the displacement threshold of the motion system, and the question arises; how does the visual system know which spot goes with which? Our answer is that the visual system doesn’t *care*. It matches the leopard’s outline (or low spatial frequencies) and the motion signal derived from this match is then spontaneously *attributed* to the spots themselves—so that the

spots appear to move with the leopard. The only disadvantage is that you could no longer see small local excursions of the spots themselves. For example, any slight ripple of the muscles in the leopard's flanks (or a change in his facial expression) would no longer be detected, but this is a small price to pay if you're trying to run away from him!

This line of reasoning is consistent with many of our observations but is also leaves several questions unanswered. For instance, if motion of the leopard's spots results from capture, why are the leaves in the background not captured as well? The only way to resolve this would be to suggest that factors such as image segmentation (Ramachandran, 1985) and segregation of figure and ground (Ramachandran and Anstis, 1986) can also directly influence the process.

Motion capture suggests a very general strategy in perception. In any dynamic visual scene the presence of a salient moving feature or frame of reference (e.g. low spatial frequencies) may dominate perception and may also cause the visual system to switch from scrutiny of individual elements to seeing them as a "texture" that belongs to the moving frame of reference. It is important to note the distinction between this model and the one proposed by Marr and Poggio (1979) and Ramachandran and Cavanagh (1985) for stereopsis. In the Marr-Poggio algorithm the low spatial frequencies are matched first and the signal from these matches is used to constrain matches made in the high-frequency domain. In our model, on the other hand, the visual system matches the low spatial frequencies (of other salient features) and then throws out matches from the finer image features. It resorts to the short cut of simply attributing the signal derived from low frequencies to all features on the object's surface.

If these ideas on motion capture are correct then perhaps the relative weights of the different masking functions depicted in Figs 2-5 are really a reflection of the statistical properties of moving bodies and their surface textures. Indeed motion capture "works" only because moving bodies normally carry their surface texture with them; spots don't normally fly off leopards. Notice, however, that the solution to the correspondence problem that we offer here is very different from three other kinds of

solutions that have been proposed in the past: (1) that establishing correspondence requires high-level cognitive processes such as "object constancy"; (2) that the perception of moving bodies involves direct "resonance" with certain invariances; and (3) that correspondence is achieved through "computation". In contrast with these three views we suggest that the system uses a set of simple tricks (such as capture) that capitalize on certain statistical properties (e.g. rigidity). By using appropriately weighted mutual inhibition between motion signals derived from different frequency bands the visual system solves correspondence and achieves continuity of object identity without involving either cognition or elaborate computation.

Acknowledgements—We thank the Natural Sciences and Engineering Research Council of Canada (grant A8606 to P.C.) for support and Ebbe Ebbesen for facilities and D. MacLeod and J. A. Deutsch for stimulating discussions.

REFERENCES

- Braddick O. J. (1974) A short-range process in apparent movement. *Vision Res.* **14**, 519-527.
- Chang J. J. and Julesz B. (1983) Displacement limits for spatial frequency filtered random-dot cinematograms in apparent motion. *Vision Res.* **12**, 1379-1386.
- Julesz B. (1971) *Foundations of Cyclopean Perception*. Univ. of Chicago Press, Chicago, Ill.
- Marr D. and Poggio T. (1979) A computational theory of human stereo vision. *Proc. R. Soc. B* **204**, 301-328.
- Marr D. (1982) *Vision*. Freeman, San Francisco, Calif.
- Nakayama K. and Silverman G. H. (1985) Detection and discrimination of sinusoidal grating displacements. *J. opt. Soc. Am.* **A2**, 267-274.
- Ramachandran V. S. (1981) Perception of apparent motion. *Studies in the Cognitive Sciences*, Vol. 11. Bulletin issued by School of Social Sciences, Univ. of California at Irvine.
- Ramachandran V. S. (1984) Entrained motion, a new visual illusion. *Invest. Ophthalm. visual Sci., Suppl.* **25**, 200.
- Ramachandran V. S. (1985) Apparent motion of subjective surfaces. *Perception* **14**, 127-134.
- Ramachandran V. S. and Anstis S. M. (1983) Displacement thresholds for coherent apparent motion in random-dot patterns. *Vision Res.* **23**, 1719-1724.
- Ramachandran V. S. and Inada V. (1984) Motion capture in random-dot patterns. *Optics News* **10**, 77.
- Ramachandran V. S. and Anstis S. M. (1986) Figure-ground segregation modulates apparent motion. *Vision Res.* **26**, 1969-1975.
- Ramachandran V. S. and Inada V. (1985) Spatial phase and frequency in motion capture of random dot patterns. *Spatial Vision* **1**, 57-67.
- Ramachandran V. S. and Cavanagh P. (1985) Subjective contours capture stereopsis. *Nature, Lond.* **317**, 527-530.