Apparent motion determined by surface layout not by disparity or three-dimensional distance

Zijiang J. He & Ken Nakayama

Vision Sciences Laboratory, Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, Massachusetts 02138, USA

THE most meaningful events ecologically, including the motion of objects, occur in relation to or on surfaces[1]. We run along the ground, cars travel on roads, balls roll across lawns, and so on. Even though there are other motions, such as flying of birds, it is likely that motion along surfaces is more frequent and more significant biologically. To examine whether events occurring in relation to surfaces have a preferred status in terms of visual representation, we asked whether the phenomenon of apparent motion would show a preference for motion attached to surfaces. We used a competitive three-dimensional motion paradigm[2] and found that there is a preference to see motion between tokens placed within the same disparity as opposed to different planes. Supporting our surface-layout hypothesis, the effect of disparity was eliminated either by slanting the tokens so that they were all seen within the same surface plane or by inserting a single slanted background surface upon which the tokens could rest. Additionally, a highly curved stereoscopic surface led to the perception of a more circuitous motion path defined by that surface, instead of the shortest path in three-dimensional space.

Using a stereoscope, observers were presented with a three-dimensional (3-D) apparent motion display that is a modification of one commonly used in 2-D displays[2]. Figure 1*a* depicts a 2-D motion display that consists of two diagonally placed tokens presented simultaneously, alternately presented with tokens located on the opposite diagonals. Such a competitive motion display will result in a bi-stable motion perception: either vertical or horizontal motion depending on the ratio of inter-token distances in the two directions. For example, when the vertical distance is kept the same, shorter horizontal distances will result in horizontally perceived motion (Fig. 1*b*), and longer distances will lead to vertically perceived motion (Fig. 1*c*). This

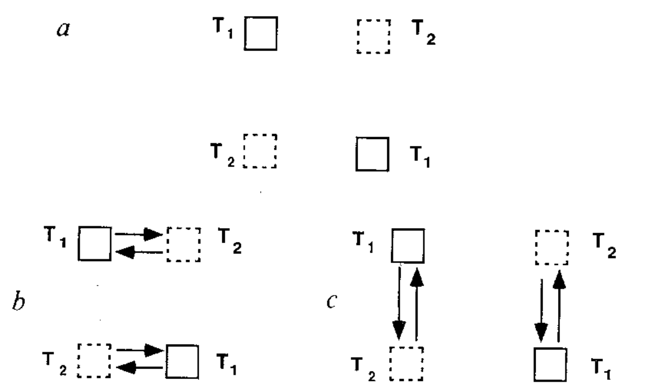
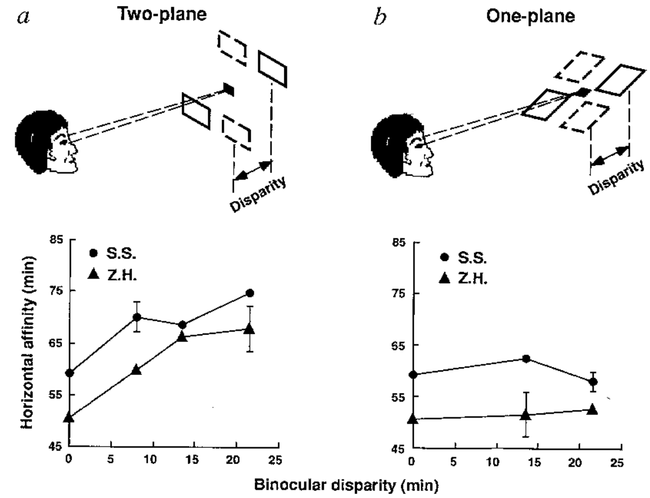


FIG. 1 Bi-stable apparent motion display where stationary targets are presented alternately at T_1 and T_2 (*a*). The relative proximity between stimuli in successive frames biases the proportion of horizontally versus vertically perceived motion. Thus horizontal motion is more likely to be seen for short horizontal distances (*b*), whereas vertical motion is more likely to be seen for longer distances (c).

'proximity' tendency[3] of the motion token to match its nearest neighbour can be summarized by determining the horizontal distance (horizontal affinity) where there is no bias for either horizontal or vertical motion. The larger the horizontal affinity, the stronger is the motion bias between the horizontal tokens. As it is a competitive apparent motion situation, a larger horizontal affinity also reflects a weaker vertical motion bias between the top and bottom tokens.

In our first experiment, the bottom and top pair of tokens had differing amounts of binocular disparity so that they could appear in the front and back vertical planes, respectively (Fig. 2*a*; two-plane condition). During the experiments, we systematically increased and decreased the horizontal distance obtaining motion judgments for each presentation. From the percentages of horizontal motion judgments for each horizontal distance, the horizontal affinity at each disparity was obtained by probit analysis[4]. Our result shows that the horizontal affinity increases with increasing disparity (Fig. 2*a*), indicating that the tendency for seeing motion within the same disparity pair of tokens increases as the disparity between the pairs of tokens increases. This result is essentially identical to those reported by Green and

FIG. 2 Comparison of results obtained from a display where increasing disparity leads to the placement of tokens in either one of two fronto-parallel planes (*a*) or where the individual tokens are slanted in depth so that they are seen to lie on a single plane, slanting back (*b*). The results indicate that as binocular disparity is increased, an increased horizontal affinity between the tokens occurs only in the two plane condition. Observers (Z.J.H., and S.S., a naive observer) viewed a computer-generated display at a distance of 100 cm through a pair of haploscopic prisms to obtain stereopsis. For each trial, the two frames of the motion display were presented alternately for 333 milliseconds for a total of six frames (three repetitions of each diagonal pair). Each token (43×21 arc min) was white (70 cd m^{-2}) and viewed against a grey (14 cd m^{-2}) background and a small black fixation square was placed midway in depth between the upper and lower pair of tokens. The observer's task was to maintain fixation and report the apparent motion direction (either horizontal or vertical) after each trial. Each block (for a given binocular disparity) consisted of 96 trials where the horizontal distances were increased and decreased repetitively. Before each session all subjects had over 100 practice trials to reduce hysteresis[2,13]. Error bars represent the mean standard error obtained for a given curve.



Odom[5], who suggested an extension of the proximity tendency to 3-D space. Thus greater 3-D distances accompanying increases in binocular disparity weaken the vertical and thus strengthen horizontal motion. Alternatively, one might consider binocular disparity itself to be a primitive feature providing a more distinctive label for determining motion direction.

Our surface-layout hypothesis, however, accounts for the result in terms of surfaces: increasing the disparity also more effectively separates the upper and lower pairs of tokens in terms of distinct non-coplanar surface planes. Horizontal affinity increases because motion within a surface plane is preferred.

To distinguish the effects of disparity or 3-D depth from our surface-layout hypothesis, we made measurements of motion perception under similar conditions of varying disparity but changed the stimulus configurations, so all tokens were perceived to lie in the same surface plane. Instead of maintaining the tokens in a fronto-parallel orientation, the small rectangular tokens were slanted stereoscopically so that all would all lie in a single imaginary plane tilted back (Fig. 2*b*). Our surface-layout hypothesis predicts less change despite increasing disparity because the top and bottom tokens still lie on a common surface plane. Using the same experimental procedure, we then compared motion perception here with the previous two-plane condition (Fig. 2*a*). The results were clear. Consistent with the surface-layout hypothesis, there is essentially no increase in the horizontal affinity despite the increasing differentiation of the top and bottom tokens in terms of binocular disparity and/or 3-D distance (Fig. 2*b*).

In the experiment shown in Fig. 2*b*, we generated a single slanted surface 'implicitly' by altering the 3-D orientation of the individual tokens so that they could be perceived as coplanar. In our next experiment we sought to generalize our findings by examining the role of a continuously perceived surface and using tokens that did not lie within the surface but which simply appeared to 'rest' on the surface. We used the original set of fronto-parallel tokens in the two-plane condition (Fig. 2*a*) and placed them in relation to two types of surface backgrounds. First, like the fronto-parallel condition of Fig. 2*a*, we used a condition where tokens were viewed against a large vertical fronto-parallel surface which had the same disparity as the farthest set of tokens (Fig. 3*a*). Second, we also used the same exact set of fronto-parellel tokens but had them rest on a receding surface defined stereoscopically as a white background with random dots (Fig. 3*b*). According to our surface-layout hypothesis, we expected that the increase in horizontal affinity with increasing disparity might be greatly attenuated, particularly in comparison to the case where there was no common plane along which the tokens could be perceived to move. Thus increasing disparity would increase the horizontal affinity only in the two-plane condition, but barely in the other. The findings shown in Fig. 3 support this view.

In addition, we have discovered a rather different phenomenon that also supports the surface-layout view, showing that the presence of a surface can radically alter the detailed trajectory of apparent motion even though the starting and ending positions of the motion tokens remain the same. We used similar motion stimuli as in the first experiment (one-plane condition), except that the four tokens were drawn to reside within an imaginary curved surface in 3-D (Fig. 4). When our subjects were presented with the tokens alone in the absence of other surfaces, the subjective motion pattern perception in the vertical direction was that of two pairs of moving rectangular tokens, resembling 'paddles',

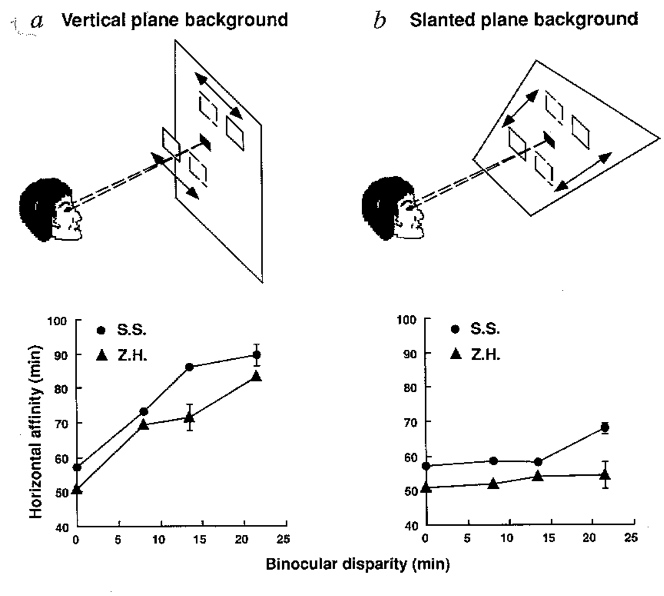
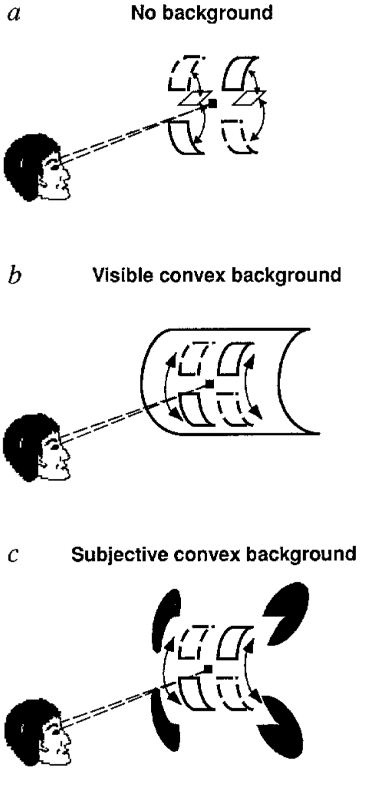


FIG. 3 The effect of visible surface backgrounds on apparent motion correspondence. Tokens here were essentially identical to the ones used in the two-plane condition of the first experiment. They were either placed against a visible vertical background (*a*), or a slanted background (*b*). The grey vertical background was rectangular (224×114 arc min), with a mean luminance of 60 cd m^{-2}, and filled with black random dots to enhance depth perception. To generate the slanted background, a similar random-dot background was differentially skewed (in each eye) to form monocular parallelograms such that when viewed stereoscopically, a receding plane was seen.

FIG. 4 Surface guided motion paths: the effect of surface background configurations on the perceived apparent motion trajectories. In each configuration, the curved and slanted tokens had exactly the same spatial configuration, constructed to lie within a semicylindrical convex surface. In *a*, the surface itself was invisible and the motion resembled paddles, flipping vertically. In *b*, this cylindrical surface is visible, defined by a grey background (60 cd m^{-2}) with black random dots having the appropriate disparity. In *c*, the cylindrical surface is subjective, defined by appropriate disparities at the corners, where it is perceived to occlude four disks. In *b* and *c*, the perceived motion path is qualitatively different from *a*, appearing to lie 'within' the cylindrical surface.



flipping across the shortest 3-D path (Fig. 4*a*). But the presence of a curved surface, defined either by a grey surface with random dots (Fig. 4*b*) or subjective contours[6] (Fig. 4*c*) led to a perceived motion path that was altogether different. Our subjects often saw sliding motion along the more circuitous path defined by the curved surface; that is, it was perceived to move 'within' the surface. We also note that the two different paths are accompanied by very different sets of correspondence rules between the edges of the tokens. With the paddle motion, the nearer and farther edges of one token correspond to the nearer and farther of the other, respectively. In the case of within-surface motion, the near edge of the token corresponds to the farther of the other and vice versa.

These experimental results taken as a whole indicate that the perception of motion is decisively affected by the perception of surfaces. Motion that is seen to be moving in relation to surfaces is favoured even if it means taking a longer path in 3-D space.

Our results argue against a simple 2-D[7,8] or 3-D[5,9] proximity rule for apparent motion. Rather than seeing motion affinity decreasing as a simple function of arbitrary 3-D distances, our work indicates the need for a representation of motion tied to surfaces. It raises the specific question as to whether a more satisfactory proximity rule might be obtained by considering distances (that is, geodesics[10,11]) within a 2-D surface embedded in 3-D space.

More generally, it raises issues regarding the nature of visual spatial representation itself. In sciences related to vision, it is often implicitly assumed that perceptual processing begins with a 2-D representation of an image followed by successive stages to form a full 3-D representation[12], both for purposes of object representation and spatially oriented behaviour. Our results provide support for an alternative view[1], arguing that the representation of surface layout is a more primitive and a more fundamental visual representation, antecedent to any abstract representation of space in terms of euclidian distances. □

Received 23 August; accepted 22 October 1993.

1. Gibson, J. J. *The Ecological Approach to Visual Perception* (Houghton Mifflin, Boston, 1979).
2. Ramachandran, V. S. & Anstis, S. M. *Nature* **304,** 529–531 (1983).
3. Burt, P. & Sperling, G. *Psychol. Rev.* **88,** 171–195 (1981).
4. Finney, D. J. *Probit Analysis* (Cambridge Univ. Press, UK, 1971).
5. Green, M. & Odom, J. V. *Science* **233,** 1427–1429 (1986).
6. Harris, J. P. & Gregory, R. L. *Perception* **2,** 235–247 (1973).
7. Ullman, S. *The Interpretation of Visual Motion* (MIT Press, Cambridge, MA, 1979).
8. Mutch, K., Smith, I. M. & Yonas, A. *Perception* **12,** 305–312 (1983).
9. Attneave, F. & Block, G. *Percept. Psychophys.* **12,** 301–307 (1973).
10. Foster, D. H. *Bio. Cyber.* **18,** 81–89 (1975).
11. Carlton, E. H. & Shepard, R. N. *J. math. Psychol.* **34,** 127–188 (1990).
12. Marr, D. *Vision* (Freeman, San Fancisco, 1982).
13. Shimojo, S. & Nakayama, K. *Perception* **19,** 285–299 (1990).

ACKNOWLEDGEMENTS. This work was supported in part by a grant from Air Force Office of Scientific Research.