

Monocular discrimination of rigidly and nonrigidly moving objects

MAARTEN A. HOGERVORST, ASTRID M. L. KAPPERS, and JAN J. KOENDERINK
Universiteit Utrecht, Utrecht, The Netherlands

We measured thresholds for the monocular discrimination of rigidly and nonrigidly moving objects defined by motion parallax. The retinal projections of rigidly moving objects are subject to certain constraints. By applying smooth 2-D transformations to the projections of rigidly moving objects, we created stimuli in which these constraints were affected. Thresholds for (generic) nonrigid transformations that in theory can be detected from rigid ones by processing pairs of views depended not only on the extent to which the rigidity constraints were affected, but also on the structure and the movement of the simulated object. Nonrigid transformations under which every three successive views had a rigid interpretation were not discriminable from rigid transformations, except in cases where the distortions were very large. Under the rigidity assumption, this would mean that a large class of nonrigidly moving objects is erroneously perceived as rigidly moving.

When moving around in the world, we have the impression that the world is stable, rigid, and three-dimensional. This is perhaps remarkable considering the fact that the retinal images change over time. Many experiments show that the visual system is capable of extracting useful information about 3-D structure from these retinal changes. The process involved is usually called structure-from-motion (SfM).

Unless assumptions are made about how the world changes over time, motion parallax does not provide us with information about depth. Without such assumptions there is a large class of possible interpretations of the underlying scene, including an interpretation formed by a set of markers moving in the frontal plane. An assumption frequently used in the literature is the *rigidity assumption*. Ullman (1979) formulated this assumption as follows: "Any set of elements undergoing a two-dimensional transformation which has a unique interpretation as a rigid body moving in space should be interpreted as such a body" (p. 146). This is a more restricted form of the processing rule proposed by Jansson and Johansson (1973) concerning the principle of minimum object change. In many experiments (e.g., those of Braunstein, Hoffman, Shapiro, Andersen, & Bennett, 1987; Johansson, 1974; Wallach & O'Connell, 1953), subjects indeed have reported perceiving a single rigid interpretation while viewing displays that have rigid interpretations. The rule generally appears to accord with reality, although exceptions are known, such as the rubber pencil illusion (Pomerantz, 1983), the Ames window (Ames, 1951), and

some of the so-called stereokinetic effects (Braunstein & Andersen, 1984; Musatti, 1924). Although these examples indicate that the rigidity assumption is not always valid, they appear to be exceptions to the general rule.

An object can also be perceived as rigidly moving in situations in which the projections are very similar to projections of a rigidly moving object. In some of the experiments of Norman and Todd (1993), subjects perceived a rigidly moving object even though they were looking at the projections of a nonrigidly moving object. We (Hogervorst, Kappers, & Koenderink, 1996) have shown that in those cases, the projections were very similar to those of a rigidly moving object. The visual system is apparently biased toward perceiving rigidly moving objects.

In modeling the visual system, one has to consider the fact that the biological hardware imposes limitations on the extraction of 3-D properties (Nakayama, 1985; Ullman, 1979). In principle two perspective projections of an object contain the information necessary to reconstruct the 3-D structure up to a uniform scaling factor (Ullman, 1979). (Note that even when more projections are provided, it is impossible to determine the uniform scaling factor from motion parallax.) Ullman (1983) showed that differences between perspective and parallel projections are small at small visual angles. However, certain 3-D properties are not defined by two parallel projections. A one-parameter family of objects exists with the same two parallel projections. These objects can be transformed into each other by applying an affine transformation consisting of scaling and shearing in depth (Koenderink & van Doorn, 1991). The 3-D properties that are not invariant under these transformations are not defined in two parallel projections and are often ill defined in two perspective projections (because in many situations parallel projection is a good approximation of perspective projection). In generic situations, two parallel projections con-

This research was supported by the Netherlands Organization for Scientific Research. Correspondence should be addressed to M. A. Hogervorst, who is now at the University of Oxford, Department of Experimental Psychology, South Parks Road, Oxford OX1 3UD, England (e-mail: maarten.hogervorst@psy.ox.ac.uk).

tain the information to determine whether or not the projections have a rigid interpretation (Bennett, Hoffman, Nicola, & Prakash, 1989; Koenderink & van Doorn, 1991; Ullman, 1977). Ullman (1979) has shown that three parallel projections of at least four points uniquely define the 3-D structure up to a reflection about the image plane (and a uniform scaling). However, whether two perspective projections or three parallel projections contain the information that the observer needs to reach a certain level of performance in a 3-D task is a matter of tolerances (Hogervorst et al., 1996; Koenderink & van Doorn, 1987).

Several authors have attempted to deduce the minimum number of views and number of elements that are required by the visual system to detect nonrigidity from parallel projections (Braunstein, Hoffman, & Pollick, 1990; Braunstein et al., 1987; Norman & Todd, 1993; Perotti, Todd, & Norman, 1996; Todd, 1982; Todd & Bressan, 1990). These experiments show that various types of nonrigidity are detectable with the theoretical minimum number of elements and views. In some experiments, performance improved with an increase in the number of views or angle of rotation (Braunstein et al., 1987; Todd, 1982). In other experiments, performance did not improve with an increase in the number of views (Todd & Bressan, 1990). Only in a few cases has the impact of the content of the views on performance been systematically analyzed. Braunstein et al. found a correlation between their measure of 3-D nonrigidity and performance in a task in which rigid and nonrigid objects had to be discriminated. It is, however, not at all clear which properties of the 2-D stimulus correlate with human performance in these tasks. Although in many situations the 3-D nonrigidity measure used by Braunstein et al. can be expected to correlate well with human performance, such a 3-D measure cannot be a good predictor in general, since different nonrigidly moving objects can have the same projections. In this paper, we relate human performance directly to the 2-D stimulus and investigate which properties of the stimulus correlate with human performance. In that sense, our experiments are more related to the experiments of Todd (1982). Todd described what characterizes the trajectories of the projections of markers forming a 3-D object that rotates with constant angular velocity about a fixed axis. He determined the impact that various changes in these characteristics have on the human ability to judge whether the objects move rigidly or not. In experiments done by Norman and Todd (1993) and Perotti et al. (1996), human perception of rigidity appears to have been determined primarily by first-order temporal relations available in two frames, although in some cases higher order temporal relations seem to have been used. In our experiments, we explore this issue further.

Many properties of egomotion and of the environmental layout are reflected in the optic flow field (e.g., time-to-contact is reflected by the divergence of the velocity field). Likewise, the optic flow field resulting from a rigidly moving set of points displays certain characteristics. Discrimination between a flow field with a rigid inter-

pretation and a flow field without a rigid interpretation comes down to finding out which of the two flow fields displays these characteristics and which one does not. In the next section, we describe what these characteristics are. Our aim is to find *the limits* for discriminating rigidly moving objects from nonrigidly moving objects from parallel projections. Such an investigation is of interest for a number of reasons. We would like to know the extent to which the visual system is biased toward perceiving a rigid world. Do we have the impression that the world is rigid because we have a realistic impression of the world, or merely because the visual system ignores nonrigidity? Also, this investigation gives insight into how the visual system combines information over time and space while performing SfM tasks. To investigate this, we used the following method. To create a stimulus without a rigid interpretation, we distorted a stimulus with a rigid interpretation—that is, we applied some 2-D transformation to the parallel projections. Then, for different types of distortions we determined the amount by which the stimulus could be distorted before it would be significantly discriminable from a stimulus with a rigid interpretation.

We investigated two types of distortions. In generic situations, two parallel projections of a nonrigidly moving object contain the information needed to determine that it moves nonrigidly. In the first experiment, we determined thresholds for distortions of this (generic) type. There are also distortions under which every pair of projections has a rigid interpretation, but in which the pairs have no common rigid solution. In the second experiment, we determined thresholds for this type of distortion.

We compared performance for objects consisting of markers randomly distributed on a *surface* with performance for objects consisting of markers randomly distributed within a *volume*. By determining thresholds for the two types of objects, we investigated whether subjects benefit from the fact that in the case of a surface, the depth is a slowly varying function of the location.

CHARACTERISTICS OF THE PROJECTIONS OF A RIGIDLY MOVING OBJECT

Koenderink and van Doorn (1991) stratified the SfM problem in such a way that one explicitly knows at which stages the various a priori assumptions are introduced and which specific geometrical expertise is required. We applied this stratification to the problem of detecting nonrigidly moving objects. We show which characteristics the projections display when they are interpretable as projections from a rigidly moving object. We use *semi-parallel projection*, by which we mean orthographic projection in which the projections are scaled with the distance between the eye and the object.

Rigid transformations—that is, translations and rotations—are a special case of affine transformations, which also include shearing and stretching transformations. If it can be shown that there is no interpretation of an object transforming in an affine way, there certainly is no rigid interpretation available. Only affine concepts (e.g.,

bisecting a line or an angle, moving lines parallel to each other) are required to detect deviations from an affine transformation. Koenderink and van Doorn (1991) showed that theoretically two parallel projections are sufficient to reveal whether an affine representation of an object exists and, if so, what it looks like: This means that the positions can be described in terms of three base vectors—the *relative* distances are known. If more than two projections are available, then for a rigid interpretation all affine representations have to be the same.

If affine methods reveal that there is an interpretation of an object undergoing an affine transformation, the transformation does not have to be a rigid one. Rigidity is a metric concept, and metric methods are required to determine whether there is a rigid interpretation.

Constraints on Every Pair of Projections

In generic situations, a pair of projections is theoretically sufficient to determine whether an object transforms rigidly or not. If two projections have a rigid interpretation, the displacement field can be made parallel by magnifying and rotating one of the projections. The reason for this is that a general rigid movement of an object can be decomposed into (1) a translation parallel to the image plane, (2) a translation orthogonal to the image plane, and (3) a rotation, which in turn can be decomposed into (3a) a rotation about an axis perpendicular to the image plane and (3b) a rotation about an axis in the image plane (Chasles's theorem). Under semiparallel projection, the only part of the movement resulting in motion parallax is rotation about an axis in the image plane. The other motion components give rise to global 2-D transformations—that is, translation (1), magnification (2), and rotation (3a). By magnifying and rotating one of the projections, motion components (2) and (3a) can be removed. The remaining motion components—(1) and (3b)—give rise to a parallel flow field. Figure 1

shows a displacement field before and after rotation and magnification. Note that in the presence of rotation and magnification, the displacement field can look very disordered. This is a constraint on every pair of projections. We refer to it as the *parallelity constraint*. If within a set of projections there is a pair whose displacement field cannot be made parallel, there is no rigid interpretation.

Additional Constraints

If the displacement field between every two projections can be made parallel, this does not mean that a rigid interpretation exists. All pairs should have a rigid interpretation in common. In general, it is difficult to see what these additional constraints are. We have shown (Hogervorst et al., 1996) what these additional constraints are for the specific case of projections resulting in a parallel flow field. Note that these constraints also apply to a set of projections that can be transformed into a parallel flow field by rotation and magnification. When there is a rigid interpretation of such a flow field, it is an object rotating about a fixed axis in the plane of projection. We use an orthogonal frame of reference in which the axis of rotation is the y -axis (lying in the image plane), the z -axis is orthogonal to the image plane, and the x -axis lies in the image plane. Figure 2 shows a view from above of a set of points rotating about an axis of rotation through one of the points parallel to the y -axis. The distances from the projected points to the projected axis of rotation (x_1, \dots, x_n) are described by the following functions:

$$\begin{cases} x_1 = r_1 \cos(\alpha_j + \phi_1) \\ \vdots \\ x_n = r_n \cos(\alpha_j + \phi_n), \end{cases} \quad (1)$$

with (r_1, \dots, r_n) the 3-D distances from the points to the axis of rotation, (ϕ_1, \dots, ϕ_n) the initial phases, and α_j the angle of rotation in the j th projection. Every projec-

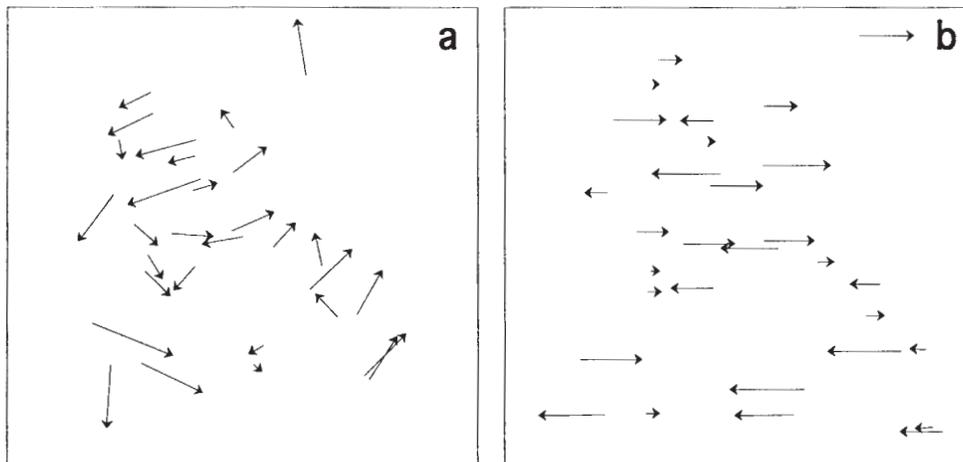


Figure 1. (a) An example of displacements between two semiparallel projections of a rigidly moving set of points. (b) By a suitably chosen rotation and magnification of one of the views, the displacement vectors can be made parallel.

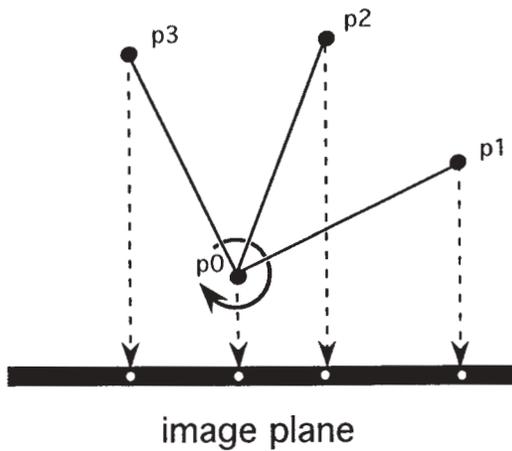


Figure 2. A view from above of a set of points rotating about a vertical axis (the y -axis) through one of the points.

tion supplies one set (x_1, \dots, x_n) of values. The 3-D structure is characterized by (r_1, \dots, r_n) and (ϕ_1, \dots, ϕ_n) . If we plot these n -vectors (x_1, \dots, x_n) in n -dimensional space with x_1 on one axis, x_2 on another, and so on, we find that all vectors lie on the same ellipse. This n -dimensional space will be referred to as *phase space*. Figure 3 shows phase space for $n = 3$. A set of points rotating rigidly about an axis in the image plane describes an ellipse in phase

space centered around the origin. It is important to note that this is a 2-D subspace: An ellipse is a planar curve. If the points do not all lie on the same ellipse in phase space, there is no rigid interpretation. This constraint will be referred to as the *elliptical path constraint*. A somewhat more relaxed constraint that one can use to detect nonrigidity is that all the points lie in a plane (the *planarity constraint*)—that is, a 2-D subspace. When all points lie in a plane, this is consistent with an object undergoing affine transformations (which leave the flow parallel). However, the transformations do not have to be rigid. Still, using the planarity constraint is often sufficient to discriminate rigid from nonrigid transformations. More details about this method can be found in Hogervorst et al. (1996).

GENERAL METHOD

Apparatus

The stimuli were generated on a Macintosh IIfx computer with a GS/C video board driving a 71-Hz Radius TDP monitor. The screen dimensions are 35.6×26.9 cm with $1,152 \times 882$ pixels.

Stimuli

We refer to the stimulus representing the rigidly (or nonrigidly) moving object as the *rigid* (or *nonrigid*) stimulus. Figure 4 shows schematically the step-by-step method used to generate the stimuli. In the first step, frames were created by orthographic projection of a set of points rotating rigidly about a vertical axis.

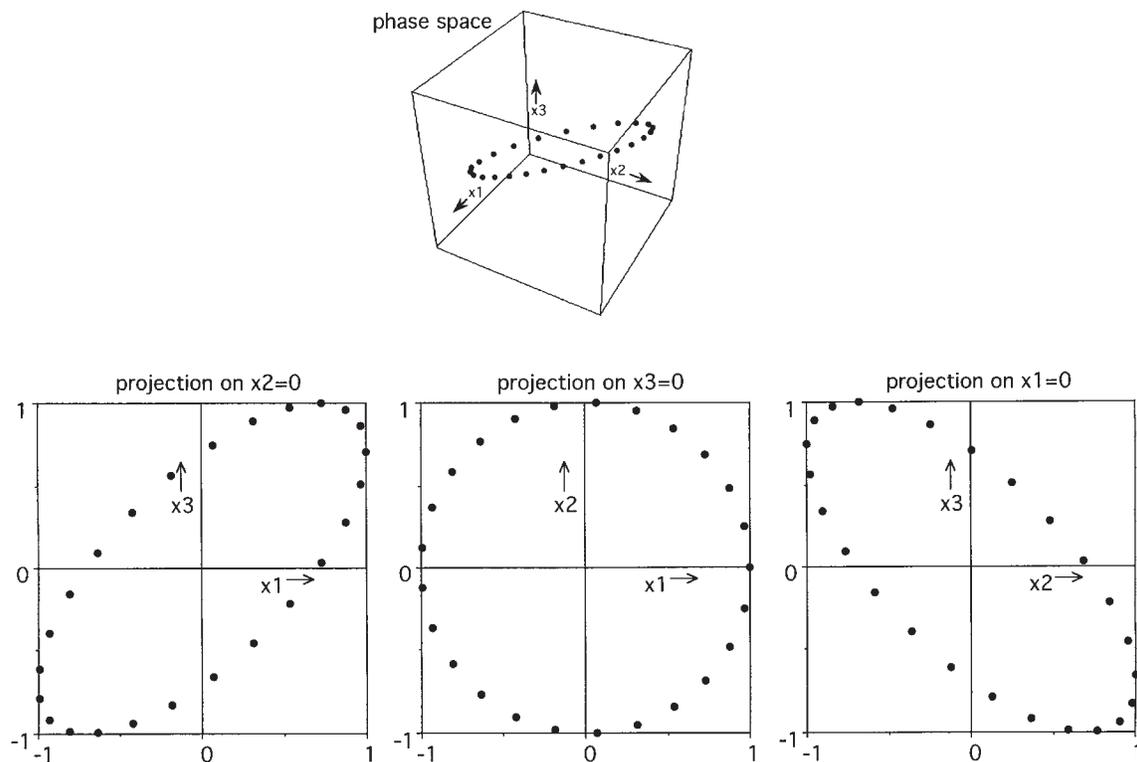


Figure 3. A phase space representation of a set of projections of three points rotating rigidly about an axis in the plane of projection. Plotted are the distances from the points to the axis of rotation in the projection. The vectors in phase space lie on an ellipse centered around the origin. Projections on different planes are also shown.

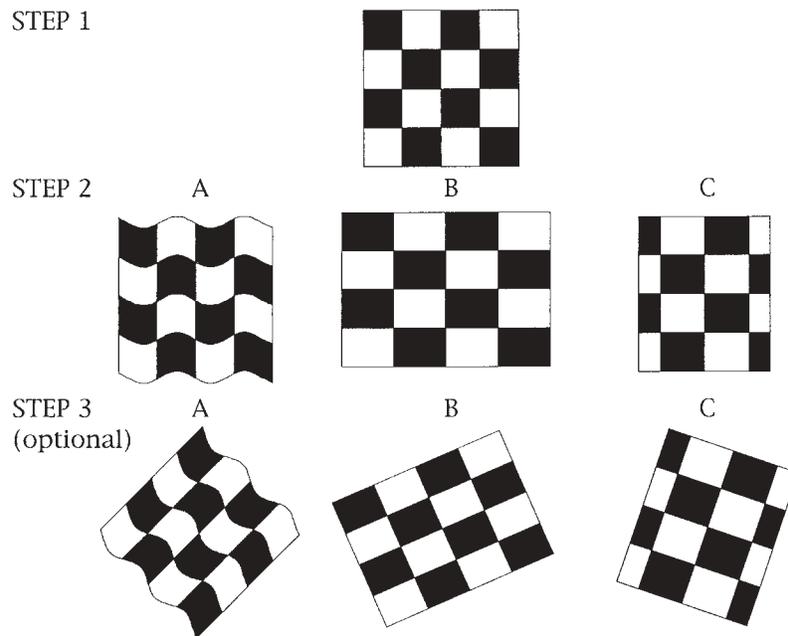


Figure 4. A schematic representation of the steps taken to create the stimuli. In Step 1, frames were created by orthographic projection of a set of points rotating rigidly about a vertical axis. To create nonrigid stimuli, these frames were transformed in Step 2. (If rigid stimuli were created, this step was skipped.) In the first experiment, we added a (time-independent) vertical offset to the position of each point that was a sinusoidally varying function of the horizontal position (transformation A). In the second experiment, the frames were stretched horizontally with a stretch that was a sinusoidally varying function of the angle of rotation. In Experiments 2A and 2B, this stretch was homogeneous (transformation B); in Experiment 2C, this was an inhomogeneous stretch—that is, the magnitude depended on the horizontal position (transformation C). In some of the experiments, the frames of the rigid and the nonrigid stimuli were subject to an additional transformation in Step 3: The frames were rotated and magnified by factors that varied over time. Specific details are given in the sections describing the experiments.

To create the nonrigid stimuli, we transformed these frames in various ways in a second step. In the first experiment, we applied transformations that affect the parallelism constraint. This was done by adding a (time-independent) vertical offset to the position of each point that was a sinusoidally varying function of the horizontal position (Figure 4, transformation A). In the second experiment, we applied transformations under which the parallelism constraint was not affected but the elliptical path constraint was affected. This was done by stretching the frames horizontally by a factor that was a sinusoidal function of the angle of rotation (of the original object about the vertical axis of rotation). In Experiments 2A and 2B, the frames were stretched homogeneously (Figure 4, transformation B). This is an affine transformation. Therefore, the task cannot be done by applying affine methods only. In Experiment 2C, the frames were stretched inhomogeneously (Figure 4, transformation C). This is not an affine transformation; that is, the affine structure changes over time. Therefore, the task can be done by applying affine methods. If the visual system mainly applies affine methods, performance should be better in Experiment 2C than in Experiment 2B.

In some of the experiments, the frames of the rigid and the nonrigid stimuli were subject to an additional transformation in a third step in which the frames were rotated and magnified by factors that varied over time. By varying the magnification from frame to frame, translation in depth could be simulated, and by varying the rotation from frame to frame, rotation about other (3-D) axes could be simulated. The nonrigid transformations were applied before the mo-

tion components were added. (If they had been applied afterward, part of the nonrigid transformation would have affected the motion instead of the structure.) Because the nonrigid transformations contain no global 2-D rotations, 2-D translations, or 2-D magnifications, they directly affect the structure. Further details are given in the sections describing the experiments.

It is possible that subjects perceive a rigidly moving object while looking at a nonrigid stimulus, but have the impression that the structure or motion of the objects represented by the nonrigid stimuli is different from that of the rigid stimuli. In that case, they would still be able to discriminate the two stimuli. To prevent subjects from using the 3-D structure or the motion to perform the task, we varied the 3-D structure from trial to trial and, except in Experiment 2A, the motion was made highly unpredictable.

Structure. The objects consisted of a number of markers, which were shown in the projections as single white pixels against a black background in high contrast. We used two types of objects: markers randomly distributed over a *surface* and markers randomly distributed within a *volume*. The objects were generated in two steps. In the first step, a standard object was generated. In the second step, an affine transformation was applied to the object to generate an object with a randomized structure but with a similar shape.

In the first step, a standard object was chosen with height and width of 6 cm (3.4° of visual angle). As standard surface, we used a vertically oriented hinged plane with a dihedral angle of 90° , except in Experiment 2C, in which we used a vertically oriented convex parabolic cylinder with a depth that was half the width. The

depth of the standard surfaces was 3 cm. The surfaces were covered with 100 markers, randomly distributed over the surface. As standard volume, we used a vertically oriented cylinder with a radius of 3 cm containing 30 markers randomly distributed within the volume. To make it easier to track the markers, the volumetric objects contained fewer markers than the surfaces. Control experiments showed no difference in performance when the number of markers was increased to 100. A different set of markers was chosen from trial to trial.

To obtain the objects used to generate the stimuli, we applied affine transformations to the standard objects. The affine transformations consisted of (1) a horizontal stretch by a factor S_x , randomly chosen between 0.8 and 1.2, (2) a stretch in depth by a factor S_z , randomly chosen between 1 and 2, and (3) a shear in depth by a factor D_z , randomly chosen between -1 and 1 :

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ D_x & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} S_x & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & S_z \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \quad (2)$$

Motion. The objects rotated about their center of mass. In Experiment 2A, a constant angular velocity was used. In the other experiments, we used what we call *semirandom movement*. In that case, the angular velocity is unpredictable, but varies smoothly over time and has a minimum and maximum. The angle of rotation between the frames [$\alpha(f)$] changes from frame to frame with a function that is the sum of three sinusoids:

$$\alpha(f) = \alpha_o \cdot \left\{ 1 + \frac{2}{3} \sum_{i=1}^3 \sin[2\pi(f \cdot \alpha_o / \lambda_i + p_i)] \right\}, \quad (3)$$

in which α_o is the average increment angle between frames, f the frame number, and λ_i and p_i are randomly chosen between 0° and 360° and between 0 and 1 , respectively. The increment angle between frames fluctuates between -1 and three times its average value. Figure 5 shows an example of such a function.

In some of the experiments, we simulated a more general movement by adding motion components in a last step (Step 3 in Figure 4). The frames were magnified and rotated, with factors changing from frame to frame. The difference in rotation and the difference in magnification changes from frame to frame in a way that is similar to the function describing the angular velocity (Equation 3), the only exception being that these properties fluctuate between -1 and $+1$ times their maximum values. This means that magnification M at frame f can be written as

$$M(f) = M(f-1) + \delta M_o \cdot \frac{1}{3} \sum_{i=1}^3 \sin[2\pi(f \cdot \alpha_o / \lambda_i + p_i)], \quad (4)$$

where $M(f-1)$ is the magnification in the previous frame and the amplitude of change in magnification is δM_o .

The rotation $R(f)$ can be written as

$$R(f) = R(f-1) + \delta R_o \cdot \frac{1}{3} \sum_{i=1}^3 \sin[2\pi(f \cdot \alpha_o / \lambda_i + p_i)], \quad (5)$$

where the amplitude of change in rotation is δR_o .

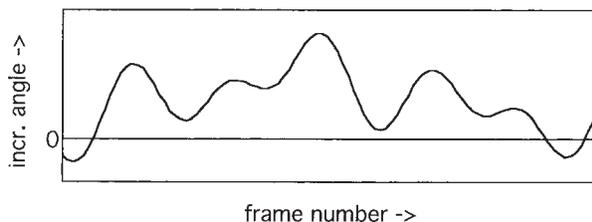


Figure 5. An example of the angular velocity as a function of time.

The angular rotation, the 2-D magnification, and the 2-D rotation functions are each defined by 6 parameters—three λ_i values and three p_i values. For each stimulus, a separate set of (6) parameters was chosen for the angular rotation, the 2-D magnification, and the 2-D rotation functions (so a total of $3 \times 6 = 18$ parameters was chosen for each stimulus).

Procedure

The stimuli were viewed monocularly from a distance of 100 cm. A rigid and a nonrigid stimulus were shown in random order. Stimulus duration was 4.2 sec. After the stimuli were shown, the subjects had to indicate which of the two stimuli was the nonrigid stimulus. With an adaptive procedure, 84% correct levels were determined for different types of nonrigid 2-D transformations. The psychometric function is modeled by an error function: The fraction of correct responses $P(A)$ as a function of the amount of nonrigidity A (the amplitude of the distortion function as used in the particular experiment) is

$$P(A) = \text{erf}(A/\sigma), \quad (6)$$

where σ is the threshold, that is, the value of A at which 84% of the responses are correct. After each trial, a maximum likelihood estimate of σ is calculated. In the next trial, we chose $A = \sigma$. The answers from 50 trials were used to calculate threshold values. It turned out that the average deviation was largely proportional to the mean. Therefore, the thresholds are presented on a logarithmic scale. The thresholds presented in this paper are weighted averages over two sessions—that is, average $y = \exp\{0.5 * [\ln(y_1) + \ln(y_2)]\}$. The errors indicated in the graphs represent the statistical errors as estimated from Monte Carlo simulations. In one session, thresholds were determined for several mixed conditions.

Subjects

Subjects M.H., S.C.P., and S.F.P. participated in the experiments. Vision was normal or corrected to normal. All subjects were experienced in psychophysical experiments. Only M.H. (the first author) was familiar with the actual way in which the stimuli were generated.

EXPERIMENT 1 Sensitivity to Deviations From the Parallelity Constraint

Method

In the first experiment, we applied nonrigid transformations that affected projections of rigidly moving objects in such a way that there was no rigid interpretation possible for every two frames. The parallelity constraint was affected by adding a vertical offset (dY) to the positions of the projected points. This offset depended on the horizontal position (x):

$$dY(x) = A \cos[2\pi(x/\lambda + p)], \quad (7)$$

where A is the amplitude of the modulation, the wavelength is λ , and phase is p . Phase p was chosen randomly between 0 and 1 from trial to trial. We determined threshold amplitudes as a function of wavelength.

Both the untransformed projections (the rigid stimuli) and the transformed projections (the nonrigid stimuli) were subject to rotation and magnification (Step 3 of Figure 4) to make the task nontrivial. Control experiments showed that when no motion components were added, thresholds were lower than the resolution of the screen and could not be measured with this setup.

We determined thresholds for two types of motion, which differed in the magnitude of the added motion components. In both situations, the amplitude of change of magnification (δM_o) was $0.36/\text{sec}$. In one case, the amplitude of the change in rotation (δR_o) was $178^\circ/\text{sec}$, and in another case it was $356^\circ/\text{sec}$. Thresholds were

obtained for wavelengths of 1.5, 3.1, 6.2, 9.3, and 12.4 cm (0.9°, 1.8°, 3.5°, 5.3°, and 7.1° of visual angle, respectively). Thresholds for both types of objects and types of motion were obtained in separate sessions.

Results

The thresholds are expressed in the maximum angle with the horizontal ($\tan \alpha = 2\pi A/\lambda$, with α being the maximum angle). Figure 6 shows the thresholds as a function of wavelength for the two types of objects and motions. The thresholds expressed as the maximal angle are largely independent of the wavelength. The fact that the threshold amplitude A of the modulation increases (almost proportionally) with increasing wavelength suggests that judgments are not based on comparisons over large distances, but are based on local comparisons, that is, on gradients in the flow field.

The thresholds (averaged over all wavelengths) ranged from 5° to 20° depending on the type of object and the motion. A formal analysis of the combined data of all subjects revealed that thresholds were significantly higher [$F(1,3) = 554.2, p < .01$] for rapidly changing rotations than for slowly changing rotations (thresholds for the high change in rotation were about 2.2 times higher on the average than those for the low change in rotation). This increase with increasing rotation could have been predicted, since when no motion components were added, the thresholds were very low. Apparently, the property of the flow reflecting the difference between the rigid and the nonrigid stimuli was masked by adding a changing rotation. This follows naturally if we assume that the low-level input consists of velocity measurements with

an uncertainty that is proportional to the speed. Koenderink and van Doorn (1987) performed a theoretical analysis on the basic limits of recovering structure and egomotion from a velocity flow field. Their study indicates that accuracy in distances decreases as (eye) rotation increases.

Thresholds for filled cylinders were significantly higher (by an average factor of 1.6) than thresholds for hinged planes [$F(1,3) = 71.1, p < .01$]. This indicates that subjects were using the fact that the flow generated from a surface is more smooth, since the depth is a slowly varying function of the location. By fitting a surface to a set of points, the noise can be reduced.

EXPERIMENT 2 Sensitivity to Deviations From the Elliptical Path Constraint

Method

In the second experiment, we transformed the projections of rigidly moving objects in such a way that every pair of frames had a rigid interpretation. To perform the task, more than two frames had to be combined. The parallelity constraint was still valid, but the elliptical path constraint was affected. The projections of a rigidly moving object were stretched horizontally by a factor that changed from projection to projection. The stretch factor $S(\alpha)$ was coupled with the angle of rotation α of the generating object:

$$S(\alpha) = 1 + A \cos[2\pi(\alpha/\lambda + p)], \quad (8)$$

where A is the amplitude of the modulation, λ is the wavelength, and p is the phase. The phase was chosen randomly between 0 and 1. Figure 7 shows a polar plot of the stretch factor as a function of angle of rotation for a wavelength of 60°. We determined threshold

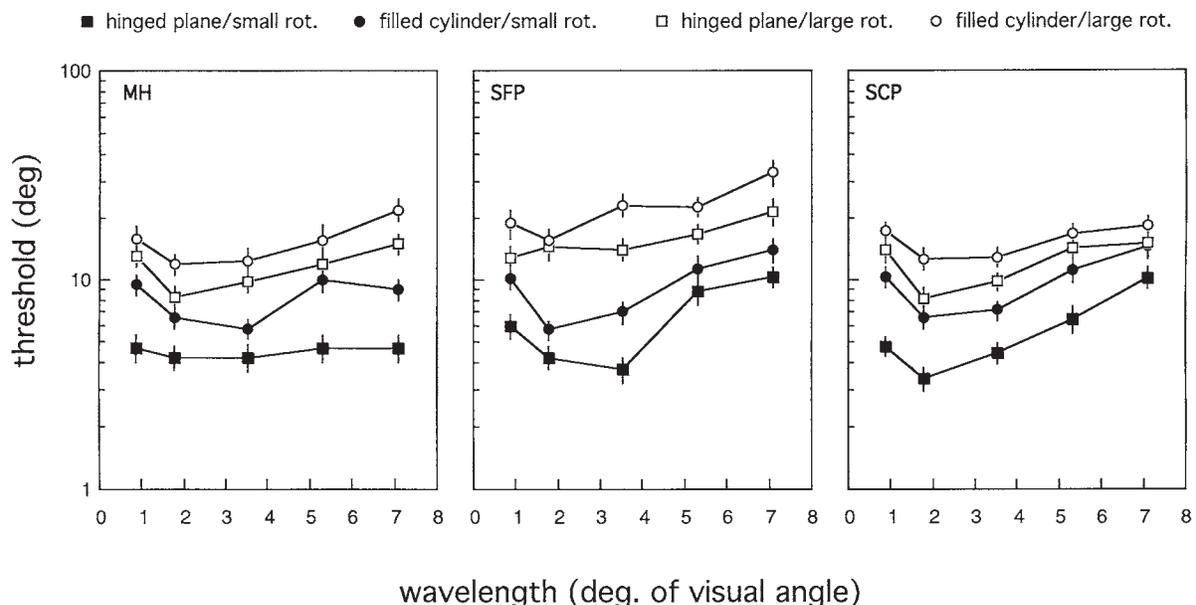


Figure 6. Thresholds for nonrigid transformations affecting the parallelity constraint (Experiment 1). The thresholds are expressed in the maximum angle that the displacement vectors make with the mean displacement direction after removal of the global rotation and magnification and are plotted as a function of wavelength (in degrees of visual angle) for two types of motion and two types of objects. Filled symbols represent large differential rotation ($\delta R_o = 356^\circ/\text{sec}$), and open symbols represent small differential rotation ($\delta R_o = 178^\circ/\text{sec}$).

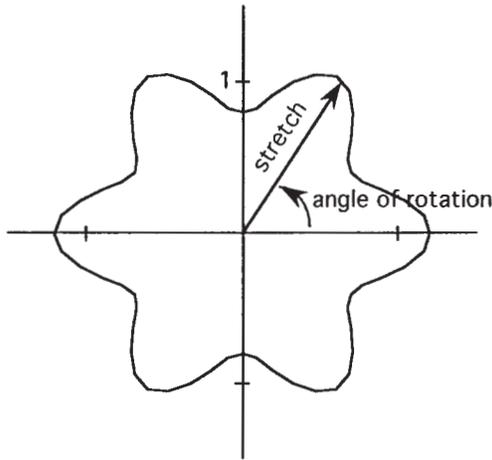


Figure 7. A polar plot of the stretch as a function of the angle of rotation for a stretching function with a wavelength of 60°.

amplitudes (A) for different wavelengths. A wavelength of 60° means, for instance, that one cycle of the stretching function corresponds to a rotation of the generating object over 60°. Because the transformation is an affine one, it is not possible to detect the nonrigidity using only affine methods.

It is important to note that the polar plot is very similar to the trajectory of the projected points in phase space. This nonrigid transformation is an affine transformation. Under an affine transformation, it is convenient to describe the vectors forming the object as linear combinations of three base vectors. We chose the following as base vectors: (0, 0, 1), (1, 0, 0), and (0, 0, 1). The transformation of the base vectors thus specifies the transformation of an arbitrary vector. Note that the third base vector (0, 0, 1) does not transform and can be omitted from the analysis. The polar plot is the same as the phase space representation of the first two base vectors rotating about the origin. Under the given nonrigid transformation, the trajectory in phase space of the object (which is an ellipse) transforms. All the vectors remain in the original plane since the transformation is affine. The transformation in the plane is the same as the transformation from a circle (with a radius of 1) to a perturbed circle in the polar plot. We compared the thresholds with thresholds pre-

dicted on the basis of a temporally second-order description of the flow field (three close views). Nonrigidity was detected when there was a set of three close views that together had no rigid interpretation. This is the case when in phase space (or in the polar plot) there is no ellipse through the three points (representing the three views) that is centered around the origin. Then the unique parabolic arc connecting the three points is curved away from the origin. The transition occurs when the curvature is zero. Figure 8 shows this transition. For small amplitudes of the stretching function (left graph), the trajectory is everywhere curved toward the origin. For a certain amplitude, there is a point with zero curvature (middle graph). For higher amplitudes, there are points at which the trajectory is curved away from the origin (right graph). At the transition point, the curvature of the trajectory in phase space at the angle with minimum stretch is zero. Using this constraint, one can derive (see Appendix) that at the transition the threshold A_{tr} is described by

$$A_{tr} = \frac{1}{1 + \left(\frac{360}{\lambda}\right)^2}, \tag{9}$$

with wavelength λ in degrees.

Experiment 2A

In this experiment, we investigated the influence of the angular velocity on the thresholds. To generate the stimuli, we started with projections of hinged planes rotating with constant angular velocity about the vertical axis. No motion components were added to the displays. Thresholds were determined for wavelengths of 30°, 90°, and 180° for angular velocities of 71°, 142°, and 284°/sec. Each session consisted of 3 (wavelengths) × 3 (angular velocities) = 9 conditions.

The aim of this experiment was to find out whether the thresholds are set by the amount of distortion of the object within a fixed time window. If so, thresholds would be lower for faster rotations than for slower rotations.

Results

Figure 9 shows the threshold amplitudes as a function of wavelength, with the angular velocity as a parameter for the 3 subjects. The thresholds increased with increas-

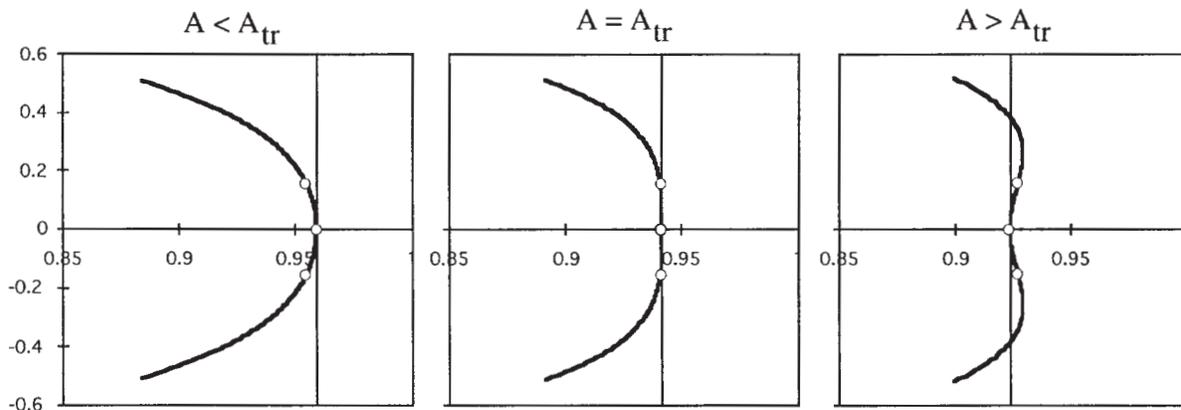


Figure 8. A polar plot of the stretching function showing the transition from a function (wavelength = 90°) with which the nonrigidity cannot be detected by processing three close views to a function with which it can be detected by processing three close views. In the left graph, the trajectory is everywhere curved toward the origin. In the middle graph, there are points for which the curvature is zero, at which the trajectory can be approximated by a straight line. In the right graph, there are points at which the trajectory is curved away from the origin. In the middle and the right cases, the nonrigidity is in principle detectable by using a second-order description of the flow.

ing wavelength. We performed a Tukey's hsd test on the combined data of all subjects. Comparing the thresholds at a certain wavelength shows that only the thresholds obtained with the highest angular velocity at a wavelength of 30° were significantly different ($p < .05$). The other thresholds were not significantly different from one another. The thresholds were largely independent of angular velocity, indicating that the subjects did not compare differences in structure over a fixed time. (If they had, the thresholds for large angular velocity would have been *lower* instead of *higher*.)

Because in this experiment we wanted to investigate the influence of the time parameter on the thresholds, constant angular velocity was used. The possibility exists that subjects perceived the nonrigid stimuli as rigid objects rotating with an angular velocity that was not constant. In that case, the stimuli could have been discriminated on the basis of perceived motion. In the next experiment, semirandom movement was used to prevent this. The pattern of results was fairly similar for both experiments, indicating that a similar strategy was used by the subjects.

Experiment 2B

In this experiment, the objects performed a semirandom movement, as described in the Method section. The angular velocity fluctuated between -1 and $+3$ times an average value (α_o) of $92^\circ/\text{sec}$. In each session, we measured thresholds for wavelengths of 30° , 90° , 120° , 150° , 180° , and 270° . Thresholds were obtained for filled cylinders and hinged planes in separate sessions.

In separate sessions, we measured thresholds for stimuli to which motion components had been added and thresholds for stimuli to which no motion components had been added. In the sessions in which we added mo-

tion components, the amplitude of the change in rotation (δR_o) was $178^\circ/\text{sec}$, and the amplitude of change in magnification (δM_o) was $0.36/\text{sec}$.

Results

Figure 10 shows the thresholds as a function of wavelength for the 3 subjects. The different curves represent different types of objects (hinged planes or filled cylinders) and different movements (with or without added motion components). The solid line shows the prediction based on the use of three close views (see Equation 9). Thresholds increased with increasing wavelength—that is, fluctuations that were fast relative to the angle of rotation were relatively easy to detect, whereas slow ones were not. The thresholds follow the prediction quite well.

A one-sample t test revealed that at all except the highest wavelength (i.e., 270°), the thresholds were significantly higher ($p < .05$) than the thresholds predicted by using three close views, for each individual subject as well as for the average over all subjects. This means that the nonrigid displays were not discriminable from the rigid ones when every set of three close views in the sequence had a rigid solution. This suggests that the subjects combined only a few views at a time; that is, they used a temporal local description of the flow. Only very large changes in structure were detected over large angles of rotation.

Let us assume that the rigidity assumption holds—that is, that stimuli that cannot be discriminated from stimuli with a rigid interpretation are perceived as rigidly moving objects. Under this assumption our results imply that—except for large distortions—nonrigid transformations under which every three successive frames have a rigid interpretation will be perceived as rigidly moving objects, although every set of three frames can have a very different rigid interpretation!

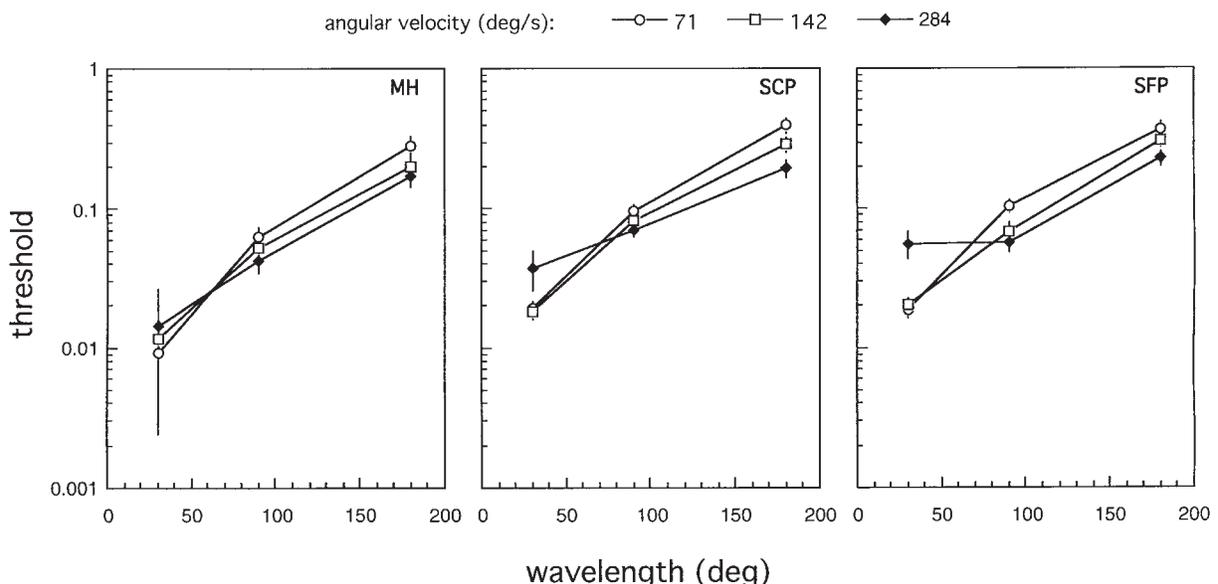


Figure 9. The threshold amplitudes for homogeneous stretching as a function of wavelength for three (constant) angular velocities for the 3 subjects (Experiment 2A).

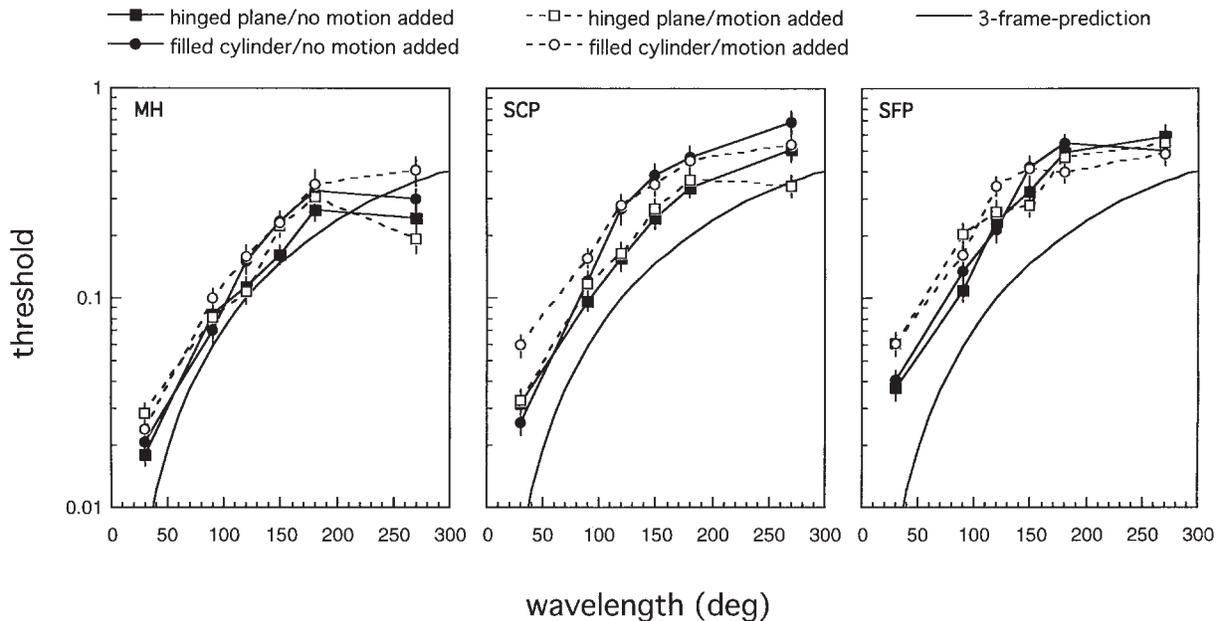


Figure 10. The threshold amplitudes for homogeneous stretching as a function of wavelength for semirandom movement for the 3 subjects. The solid line shows the prediction based on combining three close views at a time.

In other structure-from-motion experiments (e.g., those of Todd & Bressan, 1990), it has been found that subjects used a first-order temporal description of the flow field even in tasks that required a theoretical minimum of three views. Further, in the present experiment, it is possible that subjects based their judgments on a *first-order* temporal description of the flow field. However, the way the experiments were set up made it very difficult to rely on first-order information only. Any velocity field of a nonrigid stimulus would correspond to a velocity field of a rigid stimulus of a suitably chosen object and motion. Because we randomized the shape of the objects and used unpredictable motion, it would be very difficult to judge whether a given velocity field was part of a nonrigid or a rigid stimulus. In addition, the fact that performance correlated well with the prediction based on a second-order temporal description of the flow field suggests that subjects used a temporal description that was higher than first order. Although we have reason to believe that subjects did use a higher order (i.e., higher than first) description of the flow field, the possibility remains that they did not.

A formal analysis of the combined data showed a significant effect [$F(1,3) = 39.8, p < .05$] of type of object. On average, thresholds of filled cylinders were about 22% higher than thresholds for hinged planes. This indicates that—as in the first experiment—subjects were using the fact that the depth is a smooth function of location for a surface (although the difference was much smaller in this experiment than in the first experiment). This test also reveals that there was no significant effect of type of movement. The fact that the thresholds were, to a large extent, independent of the movement suggests that these results are applicable to other movements.

Experiment 2C

In this experiment, we investigated the extent to which performance might differ when the transformation is nonaffine. When the transformation is nonaffine, affine methods can be used to detect deviations from rigidity. The nonrigid transformation consisted of an inhomogeneous stretch $\tilde{S}(x, \alpha)$ that depended on the angle of rotation, as in the previous experiments, but that also depended on the horizontal position (x):

$$\tilde{S}(x, \alpha) = 1 + [S(\alpha) - 1] \cos(2\pi x/\lambda), \quad (10)$$

with $S(\alpha)$ being the stretch factor, as described by Equation 8, and with a fixed wavelength λ of 3.1 cm (= 1.8° of visual angle). If discrimination is based on the average distortion of the frames, the thresholds for the inhomogeneous stretching should be higher than for homogeneous stretching. One would expect thresholds to be lower for inhomogeneous stretching than for homogeneous stretching if subjects used the change in affine structure.

Under this nonrigid transformation, flat surfaces change into curved ones. We used vertically oriented parabolic surfaces instead of hinged planes as surfaces. This prevented subjects from indicating the curved surface as the nonrigid stimulus. Control measurements in which hinged planes were used showed no difference in results, however. We measured thresholds for surfaces and volumes in separate sessions for wavelengths of 30°, 90°, 150°, and 270°.

Results

Figure 11 shows the thresholds as a function of wavelength for the two types of objects for inhomogeneous stretching. Also presented are the thresholds for homo-

ogeneous stretching of hinged planes obtained in Experiment 2B for the situation in which no motion components were added. The solid lines represent the prediction based on the use of three close views in the case of homogeneous stretching. These lines merely function as a reference; this prediction is not valid for inhomogeneous stretching.

We performed a Tukey's hsd test on the combined data of all subjects. This showed that at each of the tested wavelengths, the thresholds for volumes were significantly higher ($p < .05$) for inhomogeneous stretching than for homogeneous stretching. This suggests that in this case, judgments were based on the average distortion. Compared with thresholds for homogeneous stretching, thresholds for surfaces under inhomogeneous stretching were significantly higher for a wavelength of 30° , not significantly different for wavelengths of 90° and 150° , and significantly lower for 270° . In a control experiment, we obtained thresholds for hinged planes that were the same as those for parabolic surfaces, suggesting that the differences were not due to differences in the type of surface. The results suggest that subjects used differences in the affine structure of surfaces over larger angles of rotation. However, subjects did not benefit from differences in the affine structure of volumes over larger angles of rotation.

DISCUSSION

We investigated *the limits* for discriminating rigidly moving objects from nonrigidly moving objects from semiparallel projections. Semiparallel projections with a rigid interpretation are subject to certain constraints. Discrimination between projections with and projections

without a rigid interpretation comes down to finding out which of both sets of projections is subject to the rigidity constraints. By distorting the projections of a rigidly moving object, we systematically affected these constraints. For different types of distortions, we obtained discrimination thresholds.

Two types of constraints have been recognized for projections with a rigid interpretation: constraints on every pair of projections (the parallelity constraint) and additional constraints on more than two projections (i.e., the elliptical path constraint in the case of parallel flow). We recognize two stages in the detection of nonrigidity. In the first stage, nonrigidity can be detected when the parallelity constraint is affected to an extent that exceeds threshold level. If it is not detected in the first stage, the nonrigidity can be detected in a second stage if additional constraints are affected to a certain extent.

In generic situations, nonrigid transformations are detectable by processing temporally first-order information or pairs of views (Ullman, 1979). Therefore, in many cases thresholds for this type of transformation determine whether the nonrigidity can be detected. This explains why many experimenters have found that two views are sufficient for the detection of nonrigidity (Braunstein et al., 1990; Braunstein et al., 1987; Petersik, 1987; Todd, 1984; Todd & Bressan, 1990).

Norman and Todd (1993) and Perotti et al. (1996) devised nonrigid stimuli that are inherently undetectable by processing temporally first-order information: Temporally higher order information is required to detect the nonrigidity. The nonrigid stimuli used by Norman and Todd consisted of objects that were subject to affine

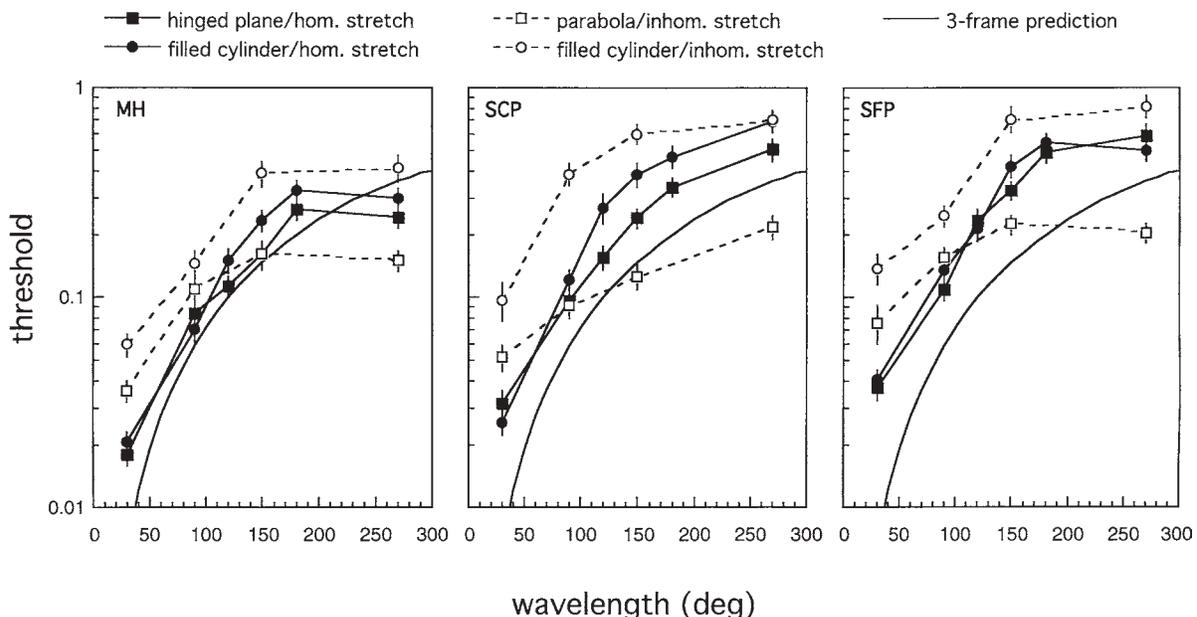


Figure 11. The threshold amplitudes for inhomogeneous stretching as a function of wavelength for the 3 subjects for two types of objects without added motion. Also shown are the thresholds of homogeneous stretching of hinged planes and filled cylinders without added motion, as obtained in Experiment 2B. The solid line shows the prediction based on combining three close views of projections under homogeneous stretching. Note that this prediction is not valid for inhomogeneous stretching.

stretching transformations while rotating. Objects that were stretched in depth were perceived as rigid, whereas objects that were stretched horizontally were perceived as nonrigid. The stimuli in these situations were very similar to the stimuli used in the present Experiment 2 with extreme amplitudes of stretch (in our paradigm)—that is, either extremely small (for stretching in depth) or extremely large (for horizontal stretches). This investigation indicated that subjects are sometimes using higher order information. Perotti et al. investigated perceived rigidity of constant flow fields (points followed flow lines of a constant velocity field). Flow fields that were subject to the parallelity constraint were perceived as rigid, although no rigid interpretation of these stimuli existed. In another experiment (similar to Todd's, 1982), they manipulated stimuli of sparse sets of points rotating about an axis in the image plane. One type of manipulations created stimuli with parallel trajectories. Some manipulations of this type were perceived as highly rigid, and others were perceived as less rigid. These studies do not clearly indicate the limits in discriminating nonrigidly moving objects from rigidly moving objects.

The difficulty of comparing our results with results from other experiments is that the task is not always the same. In some experiments, subjects have rated perceived rigidity on a rigidity scale. Notably, the constant flow fields used in Perotti et al. did not have an interpretation of a set of points moving rigidly through space, but could be interpreted as a set of points moving with constant velocity over a rigid (stationary) surface. In other experiments, subjects have had to indicate whether a stimulus was rigid or not. In our experiments, subjects had to discriminate a nonrigid stimulus from a rigid stimulus. It is not obvious that these three methods lead to the same results.

In the first experiment, we used nonrigid stimuli in which the parallelity constraint was violated. Performance depended on the structure and the movement of the simulated object. When the added motion components were larger, thresholds decreased. Apparently, the nonrigidity can be masked by adding (changing) rotation and (changing) magnification. If we assume that the low-level input to SfM consists of velocity measurements with an uncertainty that is proportional to the speed, this is understandable (Koenderink & van Doorn, 1987); for example, when rotation is added, speed increases and so does the uncertainty in the velocity measurements. In experiments of Perotti et al. (1996), it was found that perceived rigidity of constant flow fields was not affected by adding rotation. There are many differences between their and our experiments that might account for this difference in result. They did not investigate the limits. The rigidity ratings were close to 0 or 10 (on a scale between 0 and 10). Also, their task was different (rigidity rating instead of discrimination). In addition, the rotation varied over time in our experiments. Future research must address these possibilities.

Performance for points randomly positioned within a volume was higher than for points on a surface. The ve-

locity vectors in a flow field resulting from a moving surface are slowly varying over space (in direction and magnitude). The present results suggest that the internal noise in the visual system is reduced by integrating over space. This is in agreement with experiments on surface interpolation (Husain, Treue, & Anderson, 1989; Saidpour, Braunstein, & Hoffman, 1992), which indicate that the visual system can integrate motion information spatially for SfM.

Thresholds expressed in the maximum angular difference between the displacement vectors with the mean direction (after removal of global rotation and magnification) are largely independent of the wavelength. This indicates that subjects make comparisons over short distances.

In the second experiment, we determined thresholds for 2-D transformations that affect rigidity constraints on more than two projections without affecting constraints on pairs of projections. The thresholds were largely independent of the angular velocity (Experiment 2A) in the tested range. This indicates that performance does not critically depend on the time over which the object changes its structure. Detection of deviations from rigidity seems to be determined by the change in structure with respect to the change in orientation. Changes in structure that vary slowly with the angle of rotation are more difficult to detect than are fast changes. In a large range of wavelengths, the thresholds follow the prediction based on using a temporally second-order description of the flow or three close views at a time. Only for large wavelengths do deviations from this prediction occur. The predicted thresholds increase with wavelength. Only when the amplitude of the stretching function exceeds a certain level can differences in structure over larger angles of rotation be used. The results suggest that, except in the case of very large distortions, nonrigid transformations under which every three successive projections have a rigid interpretation are indistinguishable from rigid transformations.

In this experiment, performance was hardly affected by adding motion components. This indicates that the results might also be applicable to other types of movements, such as rotation about other axes.

In Experiment 2C, the nonrigid stimuli were generated by applying inhomogeneous stretching. This transformation was not affine. Therefore, the task could in principle be done by applying affine methods; the affine structure changes over time. Performance for slow distortions (i.e., for large wavelengths) was much better for surfaces under nonaffine transformation than for surfaces under affine transformation. This suggests that subjects used the change in affine structure to perform the task. This finding corresponds well with the perceived shape: For large amplitudes of distortion, subjects reported perceiving a cylindrical surface at one moment and a sinuslike surface at another moment. For objects consisting of a volume of points, performance was not better for nonaffine than for affine transformations. This suggests that subjects did not use the change in affine structure in this case. This might mean that the internal de-

scription of the object was different in these cases. A surface has a well-defined shape; for example, it changes from a parabola into a sinusform, whereas a cloud of dots remains a cloud of dots after an affine transformation.

The task we used was discriminating the nonrigid stimuli from the rigid stimuli. In principle, it is possible that subjects performed this task without perceiving rigidly and nonrigidly moving objects. However, when we asked the subjects to give their impression of the stimuli, they indicated perceiving rigidly and nonrigidly moving objects. One particularly interesting observation was that when the projections of a hinged plane were stretched homogeneously, the subjects reported perceiving a hinged plane with changing dihedral angle. This suggests that the subjects used an assumption allowing bending deformations. Subjects also reported perceiving bending deformations in the experiments performed by Jansson and Johansson (1973). Koenderink and van Doorn (1986) showed that important properties of the structure can be recovered in the presence of bending deformations.

In generic situations, discrimination is determined mainly by the extent to which the parallelity constraint is affected. In that sense, the thresholds for detecting nonrigid transformations that do not affect this constraint do not seem to be very interesting. However, they might be important for predicting *which* nonrigid transformation will be perceived. Interesting in this respect are experiments carried out by Todd (1984). In his experiments, subjects had to judge the curvature of cylinders rotating about a horizontal axis. Performance did not deteriorate when horizontal fluctuations were introduced to the (vertical) trajectories of the projected points. This suggests that the nonrigidity is limited to the horizontal direction. This stimulus is somewhat similar to the stimuli used in our first experiment. In this experiment, subjects reported perceiving rigidly moving objects with dots moving over the surface. One can hypothesize that a similar percept arises when the objects over which the dots appear to move undergo nonrigid transformations that do not affect the parallelity constraint and affect the additional constraints below threshold level.

In conclusion, the results indicate that comparisons in structure are mainly made locally in time and in space. Nonrigid transformations that cannot be detected by using a temporal second-order description of the optic flow can be detected as nonrigid only when the changes in structure are very large. This means that there must be a large class of nonrigid transformations that cannot be detected by the visual system. If we assume that the rigidity assumption holds, this means that a large class of nonrigidly moving objects will be erroneously perceived as rigidly moving.

REFERENCES

- AMES, A. (1951). Visual perception and the rotating trapezoidal window. *Psychological Monographs*, **65** (7, No. 324).
- BENNETT, B. M., HOFFMAN, D. D., NICOLA, J. E., & PRAKASH, C. (1989). Structure from two orthographic views of rigid motion. *Journal of the Optical Society of America A*, **6**, 1052-1069.
- BRAUNSTEIN, M. L., & ANDERSEN, G. J. (1984). A counterexample to

- the rigidity assumption in the visual perception of structure from motion. *Perception*, **13**, 213-217.
- BRAUNSTEIN, M. L., HOFFMAN, D. D., & POLLICK, F. E. (1990). Discriminating rigid from nonrigid motion: Minimum points and views. *Perception & Psychophysics*, **47**, 205-214.
- BRAUNSTEIN, M. L., HOFFMAN, D. D., SHAPIRO, L. R., ANDERSEN, G. J., & BENNETT, B. M. (1987). Minimum points and views for the recovery of three-dimensional structure. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 335-343.
- HOGERVORST, M. A., KAPPERS, A. M. L., & KOENDERINK, J. J. (1996). Structure from motion: A tolerance analysis. *Perception & Psychophysics*, **58**, 449-459.
- HUSAIN, M., TREUE, S., & ANDERSEN, R. A. (1989). Surface interpolation in three-dimensional structure-from-motion perception. *Neural Computation*, **1**, 324-333.
- JANSSON, G., & JOHANSSON, G. (1973). Visual perception of bending motion. *Perception*, **2**, 321-326.
- JOHANSSON, G. (1974). Visual perception of rotary motion as transformation of conic sections. *Psychologia*, **17**, 226-237.
- KOENDERINK, J. J., & VAN DOORN, A. J. (1986). Depth and shape from differential perspective in the presence of bending deformations. *Journal of the Optical Society of America A*, **3**, 242-249.
- KOENDERINK, J. J., & VAN DOORN, A. J. (1987). Facts on optic flow. *Biological Cybernetics*, **56**, 247-254.
- KOENDERINK, J. J., & VAN DOORN, A. J. (1991). Affine structure from motion. *Journal of the Optical Society of America A*, **8**, 377-385.
- MUSATTI, C. L. (1924). Sui fenomeni stereocinetici [On stereokinetic phenomena]. *Archivio Italiano di Psicologia*, **3**, 105-120.
- NAKAYAMA, K. (1985). Higher order derivatives of the optical velocity vector field: Limitations imposed by biological hardware. In D. Ingle, M. Jeannerod, & D. Lee (Eds.), *Brain mechanisms and spatial vision* (pp. 59-71). Dordrecht: Martinus Nijhoff.
- NORMAN, J. F., & TODD, J. T. (1993). The perceptual analysis of structure from motion for rotating objects undergoing affine stretching transformations. *Perception & Psychophysics*, **53**, 279-291.
- PEROTTI, V. J., TODD, J. T., & NORMAN, J. F. (1996). The visual perception of rigid motion from constant flow fields. *Perception & Psychophysics*, **58**, 666-679.
- PETERSIK, J. T. (1987). Recovery of structure from motion: Implications for a performance theory based on the structure-from-motion theorem. *Perception & Psychophysics*, **42**, 355-364.
- POMERANTZ, J. R. (1983). The rubber pencil illusion. *Perception & Psychophysics*, **33**, 365-368.
- SAIDPOUR, A., BRAUNSTEIN, M. L., & HOFFMAN, D. D. (1992). Interpolation in structure from motion. *Perception & Psychophysics*, **51**, 105-117.
- TODD, J. T. (1982). Visual information about rigid and nonrigid motion: A geometric analysis. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 238-252.
- TODD, J. T. (1984). The perception of three-dimensional structure from rigid and nonrigid motion. *Perception & Psychophysics*, **36**, 97-103.
- TODD, J. T., & BRESSAN, P. (1990). The perception of 3-dimensional affine structure from minimal apparent motion sequences. *Perception & Psychophysics*, **48**, 419-430.
- ULLMAN, S. (1977). *The interpretation of visual motion*. Unpublished doctoral dissertation, MIT.
- ULLMAN, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.
- ULLMAN, S. (1983). Recent computational studies in the interpretation of structure from motion. In A. Rosenfeld & J. Beck (Eds.), *Human and machine vision* (pp. 459-480). New York: Academic Press.
- WALLACH, H., & O'CONNELL, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, **45**, 205-217.

APPENDIX

In this appendix, we derive for the case of homogeneous stretching the theoretical limits for detecting nonrigidity from three close views—that is, a second-order temporal description of the optical flow field.

The nonrigidity is detectable when there is a set of three close views that have no rigid interpretation. This is the case when there is a set of three close views in which the points in phase space cannot be connected by an ellipse centered around the origin. Then the *unique* parabolic arc connecting the three points is curved away from the origin. The transition occurs when the curvature is zero. For the type of nonrigidity used in Experiments 2A and 2B (homogeneous stretching), the trajectory in phase space is curved away most at the angle with minimum stretch. Therefore, the nonrigidity is detectable when the curvature of the trajectory in phase space at the angle with minimum stretch is zero. The trajectory can be approximated locally (i.e., using a second-order Taylor expansion) by a straight line. The stretch function is given by

$$S = 1 - A \cos(2\pi\alpha/\lambda), \quad (\text{A1})$$

where S is the stretch as a function of the angle of rotation α , A is the amplitude, and λ is the wavelength of the modulation function. This function can be locally approximated by

$$\tilde{S} = 1 - A(1 - 2\pi^2/\lambda^2 \alpha^2). \quad (\text{A2})$$

The equation of a straight line ($x = 1 - A$) in polar coordinates is

$$R = (1 - A)/\cos(\alpha), \quad (\text{A3})$$

which can be approximated by

$$\tilde{R} = (1 - A) \cdot (1 + \frac{1}{2}\alpha^2). \quad (\text{A4})$$

The transition occurs when both approximations are equal, $\tilde{S} = \tilde{R}$, which is the case when

$$A = \frac{1}{1 + \left(\frac{2\pi}{\lambda}\right)^2}, \quad (\text{A5})$$

or

$$A = \frac{1}{1 + \left(\frac{360}{\lambda}\right)^2}, \quad (\text{A6})$$

when the wavelength is expressed in degrees instead of radians.

(Manuscript received December 12, 1995;
revision accepted for publication December 17, 1996.)