MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 721 June, 1983

MAXIMIZING RIGIDITY: THE INCREMENTAL RECOVERY OF 3-D STRUCTURE FROM RIGID AND RUBBERY MOTION

Shimon Ullman

ABSTRACT:

The human visual system can extract 3-D shape information of unfamiliar moving objects from their projected transformations. Computational studies of this capacity have established that 3-D shape, can be extracted correctly from a brief presentation, provided that the moving objects are rigid. The human visual system requires a longer temporal extension, but it can cope, however, with considerable deviations from rigidity. It is shown how the 3-D structure of rigid and non-rigid objects can be recovered by maintaining an internal model of the viewed object and modifying it at each instant by the minimal non-rigid change that is sufficient to account for the observed transformation. The results of applying this incremental rigidity scheme to rigid and non-rigid objects in motion are described and compared with human perceptions.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support is provided by National Science Foundation Grant 79-23110MCS.

1. The rigidity–based recovery of structure from motion

1.1 The perception of structure from motion by human observers

The human visual system is capable of extracting three–dimensional (3–D) shape information from two-dimensional (2–D) transformations in the image. Experiments employing shadow projections of moving objects and computer generated displays have established that the 3–D shape of objects in motion can be perceived when their changing projection is observed, even when each static view is completely devoid of 3–D information.

Observations related to this intriguing capacity have been reported as early as 1860 by Sinsteden (described in Miles 1931), who observed the perception of depth and depth–reversals induced by a distant windmill, and then examined this phenomenon in the laboratory using rotating cardboard objects. The early experiments in this area where concerned primarily with the perceived depth–reversals of rotating objects. The fact that the 3–D structure of moving objects can be recovered perceptually from their changing projection was noted by Musatti in 1924 (see Johansson 1978). It was investigated systematically for the first time using shadow projections by Wallach & O'Connell (1953) who coined the term "kinetic depth effect" to describe this phenomenon. The perception of structure from motion has been investigated extensively since, under various conditions, including the motion of connected and unconnected elements, and using both perspective and orthographic projections. (In orthographic projection the projecting rays are parallel, and perpendicular to the image plane; in a perspective projection they meet at a common point.) For detailed reviews see (Braunstein 1976, Johansson 1978, Ullman 1979a).

1.2 Computational studies of the recovery of structure from motion

In trying to recover 3–D structure from the transformations in the image, one is faced with the problem of inherent ambiguity: unless some constraints are imposed, the image transformations are insufficient to determine the 3–D structure uniquely. This ambiguity problem has been the focus of a number of computational studies that attempted to discover the conditions under which 3–D structure is uniquely determined by the projected transformations in the image.

From the earliest empirical studies of the kinetic depth effect it has been suggested that the rigidity of objects may play a key role in the perception of structure from motion (Wallach & O'Connell 1953, Gibson & Gibson 1957, Green 1961, Johansson 1975). Computational studies have established that rigidity is a sufficiently powerful constraint for imposing uniqueness upon the 3–D interpretation of the viewed transformations. A 2–D transformation can be tested to determine whether or not it is compatible with the projection of a rigid object in motion. If it is, then the inducing object is in general guaranteed to be unique, and its 3–D structure can be recovered. (Under orthographic projection the structure is determined uniquely up to a reflection about the image plane. This is an unavoidable ambiguity, since the orthographic projections of a rotating object, and its mirror image rotating in the opposite direction, coincide.)

Uniqueness results have been established under a number of different conditions. Ullman & Fremlin (Ullman 1979a) have shown that under orthographic projection three views of four non-coplanar points are sufficient to guarantee a unique 3–D solution. Longuet-Higgins & Prazdny (1980) proved that the instantaneous velocity field and its first and second spatial derivatives at a point admit at most three different 3–D interpretations. Recently, Tsai & Huang (1982) in an elegant analysis have shown that, with the exception of a few special configurations, two perspective views of seven points are also sufficient to guarantee uniqueness. Additional uniqueness results have been obtained for restricted motion, such as planar surfaces in motion (Hay 1966), the recovery of time to collision (Lee 1976,1980), pure translatory motion (Clocksin 1980), planar or fixed axis motion (Bobick, 1983; Hoffman & Flinchbaugh, 1982; Webb & Aggarwal, 1981) and translation perpendicular to the rotation axis (Longuet-Higgins 1983,). A review of these and other computational results obtained to date on the recovery of structure from motion will appear in (Ullman 1983).

In summary, the uniqueness results establish that by exploiting a rigidity constraint the recovery of 3–D structure is possible on the basis of motion information alone, and that the recovery is possible in principle on the basis of information that is local in space and time.

1.3 Additional requirements for the recovery of structure from motion

There are a number of interesting differences between the mathematical results cited above and the recovery of structure from motion by the human visual system.

Extension in time: Although the recovery of structure from motion is possible in principle from the instantaneous velocity field, the human visual system requires an extended time period to reach an accurate perception of the 3–D structure (Wallach & O'Connell 1953, White & Mueser 1960, Green 1961). This difference is not surprising, especially when the recovery scheme is applied locally to small objects or local surface patches. For surface patches extending about 2^0 of visual angle, drastically different objects can induce almost identical instantaneous velocity fields. This limitation of the instantaneous velocity field is illustrated in figure 1. The figure shows a cross-section of two surfaces, $S1$ and $S2$ seen from a side view. The surfaces are assumed to be rotationally symmetric around the observer's line of sight, so that $S1$, for example, is a part of the surface of a sphere. When the viewing distance is such that the surfaces in Figure 1 occupy 2^0 of visual angle, the difference in their velocity fields within the entire patch does not exceed 6%. The implication is that although the instantaneous velocity field contains sufficient information for the recovery of the 3–D shape, the reliable interpretation of local structure from motion requires the integration of information over a more extended time period.

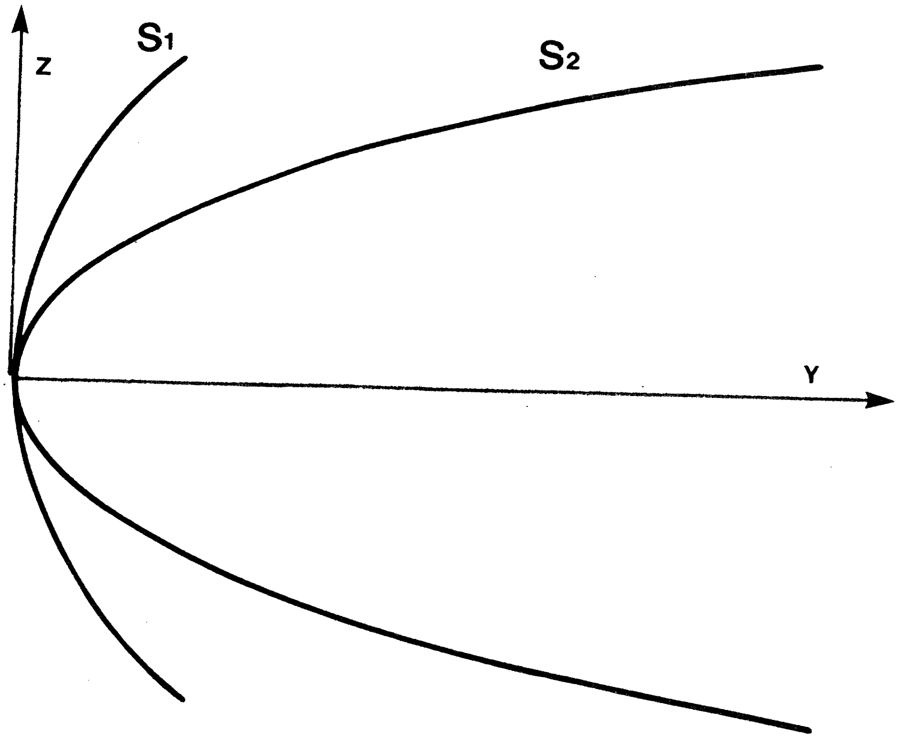

Figure 1 Limitations of the instantaneous velocity field. S_1 and S_2 are two rotationally symmetric surfaces when the surfaces occupy 2^0 of visual angles, the differences in their velocity fields at each point does not exceed 6%.

Deviations from rigidity: In interpreting structure from motion, the visual system can cope with less than strict rigidity (Johansson 1964, 1978, Jansson & Johansson 1973). If the viewed object undergoes a rigid transformation combined with some non-rigid distortions, the 3-D object and its distortions can often be perceived. This tolerance for deviations from rigidity also implies that the recovery process enjoys a certain immunity to noise. If the viewed object is in fact rigid, but the measurement of its motion and the computations performed are not entirely accurate, the result would not be a complete breakdown of the recovery process, but a perception of a slightly distorting object. Robustness with respect to errors in the measured velocity field is particularly important, since an accurate measurement of the velocity field is difficult to obtain (Fennema & Thompson 1979, Horn & Schunck 1981, Adelson & Movshon 1982, Ullman & Hildreth 1983). For human observers, kinetic depth displays of rigid objects often give rise to the perception of a somewhat distorting object (Wallach, Weisz & Adams, 1956; White & Mueser 1960, Green 1961, Braunstein 1962).

Successive approximation: For the human visual system, the recovery of structure from motion is not an all–or–none process. Although the accuracy of the perceived structure improves with time, a cruder perception of shape and relative depth are still possible under short viewing periods. For short viewing times, objects often appear flatter than their correct 3–D structure. For example, a rotating cylinder (Ullman 1979a) would not appear to have the full depth of the correct 3–D object. The perceived shape is qualitatively similar to the correct one, and, as mentioned above, often improves with time.

Integrating sources of 3–D information: The kinetic depth experiments demonstrated the capacity of the human visual system to recover 3–D structure on the basis of motion information alone. Wallach & O'Connell, for example, included in their experiments wireframe objects whose static projection induced no 3–D perception. Such objects appear initially as flat and lying in the frontal image plane, but acquire the correct 3–D shape when viewed in motion. Under more natural conditions, motion information is seldom dissociated from other sources of information. In this case, problems related to the integration of different sources of information arise. Wallach, O'Connell & Neisser (1953), for instance, reported a kinetic depth experiment with an object whose static projection was often perceived as a 90^0 corner (i.e., three mutually perpendicular rods). The actual shape of this object was different: the angle between the rods was in fact 110 rather than 90^0. Observers who initially saw the 90^0 corner, often perceived a 3-D structure that agreed better with the object's correct structure after viewing the projection of the moving object for sufficiently long time. The initially perceived structure is therefore not necessarily flat, and it may be influenced by different sources of 3–D information. In the above experiment the structure from motion interpretation sometimes eventually dominated over the static cues. There are also cases of conflicting 3–D information in which the perceived 3–D structure is determined primarily by the static cues rather than the motion information (Ullman 1979a, ch. 5). The case where the structure from motion process is ineffective is not directly relevant to the present discussion, and will not be considered further.

1.4 A hypothesis: maximizing rigidity relative to the current internal model

The above discussion suggests that to be comparable in performance to the human visual system the process of recovering structure from motion should meet the following requirements. (i) At each instant there should exist an estimation of the 3–D structure of the viewed object. This internal model of the viewed structure may be initially crude and inaccurate, and may be influenced by static sources of 3–D information. (ii) The recovery process should prefer rigid transformations. (iii) The recovery scheme should tolerate deviations from rigidity. (iv) It should be able to integrate information from an extended viewing period. (v) It should eventually recover the correct 3–D structure, or a close approximation to it.

Most of these requirements can be met naturally by the following "incremental rigidity" scheme. Assume that at any given time there is an internal model of the viewed object. Let $M(t)$ denote the internal model at time t. As the object continues to move, its projection would change. If $M(t)$ is not an accurate model of the object at time t, then no rigid transformation of $M(t)$ would be sufficient to account for the observed transformation in the image. The crucial step is that the internal model would then be modified by the minimal change that is still sufficient to account for the observed transformation. In other words, the internal model resists changes as much as possible, and consequently becomes as rigid as possible.

Such a scheme takes into account the use of a current 3–D model which initially may be inaccurate (requirement i), the tendency to perceive rigid transformations when possible (requirement ii), without requiring strict rigidity (requirement iii). It also combines information from extended viewing periods, by incorporating incremental changes into the internal model (requirement iv). It thus has the appealing property of combining information from extended periods, and at the same time using at any instant only the internal model and the incoming image at that instant. It does not require storing and using long sequences of different views of the object as might be used, for example, in a computer implementation of a structure from motion process. Unlike Johansson's (1974) trajectory–based scheme, which also integrates information over an extended viewing period, this mode of temporal integration is not limited to fixed–axis motion, but can be applied to objects under general motion.

The incremental rigidity scheme therefore meets four of the five requirements listed above. It remains unclear, however, whether such a scheme can cope with the last, and most crucial, requirement. That is, if $M(t)$ is initially incorrect, would it eventually converge to the correct 3–D structure? The answer is not obvious: if the model is incorrect at time t, it is not clear whether an attempt to transform it as rigidly as possible would bring it any closer to the unknown structure of the viewed object. To assess the feasibility of the incremental rigidity scheme it is therefore necessary to examine whether the incremental changes in $M(t)$ would cause it in the long run to converge to the correct 3–D structure of the viewed object. The main problem regarding the incremental rigidity scheme is therefore the following: if $M(t)$ is updated by transforming it at each instant as rigidly as possible, will it converge eventually to the correct 3–D structure under rigid motion, and under deviations from rigidity. This problem is examined in the next two sections. Section 2 describes more precisely the incremental rigidity scheme and how it is applied to recover 3–D structure from motion. Section 3 describes the results of applying this scheme to rigid as well as non–rigid objects. The general conclusion is that the incremental rigidity scheme copes successfully with rigid objects as well as with considerable deviations from rigidity, and that it resembles various aspects of the perceptual recovery of structure from motion.

2. The incremental rigidity scheme

Analytic treatment of the convergence requirement did not seem tractable; therefore, a computer program was devised to test the convergence of the incremental rigidity scheme to the correct 3–D structure under both rigid and non–rigid motion. This section will describe the incremental rigidity scheme that has been employed. Section 2.1 describes the basic scheme, 2.2 considers possible modifications, and 2.3 examines briefly problems of efficiency, including the execution of the scheme in an analogue, distributed manner. Section 3 then describes the behavior of the scheme as revealed by the computer simulations, and in particular its convergence to the correct 3–D structure under rigid and non–rigid motion.

2.1 The basic scheme

For the computer implementation it is convenient to consider the visual input as a sequence of frames, each one depicting a number of identifiable feature points rather than continuous flow. (The temporal discreteness of the input is not a necessary aspect of the scheme, a continuous formulation is also conceivable.) The scheme maintains and updates an internal model $M(t)$ of the viewed object. $M(t)$ consists of a set of three–dimensional coordinates (X_i, Y_i, Z_i). Assuming orthographic projection onto the $X - Z$ image plane, (X_i, Z_i) are the image coordinates of the i^{th} point, and Y_i is its depth as estimated in the current model. ($X - Z$ was chosen as the image plane, with the positive Y direction pointing away from the observer. This notation keeps the coordinate system right–handed.) As will be noted below, a similar scheme can be defined for a perspective rather than orthographic projection. For small objects, or small surface patches of objects, the two projections are in close agreement, hence, the type of projection employed has little effect on the behavior of the scheme. The relation between the two projections is discussed in more detail in section 4.1. In the lack of information about the 3–D shape of the viewed object, the initial model $M(t)$ at $t = 0$ is taken to be completely flat, i.e., $Y_i = 0$ (or any other constant, since the overall distance to the object remains undetermined) for $i = 1 \ldots n$, where n is the number of points considered in the computation.

Next, a new frame corresponding to a later time t' is considered, and the problem is then to update $M(t)$ so as to agree with the new frame, while making the transformation from $M(t)$ to $M(t')$ as rigid as possible. The new frame is represented as a set of 2–D image coordinates (x_i, z_i). The depth values y_i are as yet undetermined. It is assumed, however, that the correspondence between points in the two successive frames is known. When the y_i values have been estimated, the set of coordinates (x_i, y_i, z_i) is the estimated structure at time t', denoted by $S(t')$. The notation convention used is that all the parameters that refer to $M(t)$ are denoted by capital letters (X, Y, Z, etc.), and those referring to $S(t)$ by small letters (x, y, z, etc.).

The most rigid transformation of the internal model $M(t)$ is now determined in the following manner. Let L_{ij} denote the distance between points i and j in $M(t)$. That is:

$$L_{ij}^2 = (X_i - X_j)^2 + (Y_i - Y_j)^2 + (Z_i - Z_j)^2 \tag{1}$$

Similarly, l_{ij} is the internal distance in the estimated structure between points i and j at time t'. That is:

$$l_{ij}^2 = (x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2 \tag{2}$$

A rigid transformation implies that $L_{ij} = l_{ij}$ for all i, j (that is, all the internal distances in the object remain unchanged). To make the transformation as rigid as possible, the unknown depth values y_i should therefore be chosen so as to make the values of l_{ij} and L_{ij} agree as closely as possible. If $D(L_{ij}, l_{ij})$ is a measure of the difference between L_{ij} and l_{ij}, then the problem of determining the most rigid transformation of the model can be formulated as determining the values of y_i so as to minimize the overall deviation from rigidity $\sum_{i,j} D(L_{ij}, l_{ij})$ $(i = 1, \ldots, n-1, j = i+1, \ldots, n)$.

A reasonable choice of the distance function D should make the contributions from nearby points weigh more than distant ones. The reason is that the nearest neighbors to a given point are more likely to belong to the same object than distant neighbors. A point is consequently more likely to move rigidly with its nearest neighbors. An example of such a distance measure is:

$$D(L_{ij}, l_{ij}) = \frac{(L_{ij} - l_{ij})^2}{L_{ij}^3} \tag{3}$$

In this measure the effect of, say, a 10% change in L_{ij} decreases as L_{ij} increases.

After the values of y_i have been determined using the minimization criterion, (x_i, y_i, z_i) becomes the new model $M(t')$. A new frame is then registered, and the process repeats itself.

In summary, the computation involved at each step in establishing the most rigid interpretation is the following. Given an internal model $M(t)$ in the form $(X_i, Y_i, Z_i), i = 1, \ldots n$, and the new frame $(x_i, z_i), i = 1, \ldots, n$, find a vector of depth-values y_i such that the overall deviation from rigidity $\sum_{i,j} D(L_{ij}, l_{ij})$ for $i = 1, \ldots, n-1, j = i+1, \ldots, n$ is minimized.

2.2 Possible modifications

Some modifications of the basic scheme presented above are possible. For example, a somewhat different form of the metric D can be used. The important issue to explore, however, is whether any such scheme converges successfully to the correct 3–D structure. As discussed above, the incremental rigidity scheme meets requirements (i) through (iv), but it is unclear whether it can also meet the convergence requirement. To be considered a plausible scheme for the recovery of structure from motion by the human visual system, the convergence requirement must also be met for rigid as well as not strictly rigid objects. If a particular version of the scheme accomplishes the 3–D recovery task successfully, then it provides a certain existence proof that an incremental rigidity scheme can meet all of the requirements listed above.

Two modifications of the basic scheme described above were explored. One was to introduce some changes to the metric D. The other, more substantial modification, takes into account the fact that $M(t)$, the internal model at time t, may be inaccurate, and allows it to be corrected.

The basic scheme described in the previous section can be summarized as minimizing $D(M(t), S(t'))$, a measure of the overall distortion between the 3–D model at time t and the new computed structure at time t'. The modified method searches for a modified, corrected model $M'(t)$, such that the transition from $M(t)$ to $M'(t)$ (the correction to the internal model) is small, and the transition from $M'(t)$ to $S(t')$ is as rigid as possible. This modified scheme minimizes therefore the sum: $D(M(t), M'(t)) + D(M'(t), S(t'))$. Since it allows changes in the internal model $M(t)$, this scheme will be referred to below as the "flexible model" scheme. In general, the modifications explored of the metric D had only small effects on the convergence to the correct 3–D structure. The use of the more complicated flexible model scheme also did not introduce fundamental changes, but usually resulted in an overall improvement of the computed structure. This flexible model also has the advantage that other 3D cues could influence the transition from $M(t)$ to $M'(t)$ (like stereo or shading cues that change dynamically). These observations suggest that the basic incremental rigidity scheme is not sensitive to variations in the exact formulation of the minimization problem. Additional comments regarding the modified scheme are incorporated in the discussion of the results in section 3.

2.3 Implementation

The incremental rigidity scheme described above has been implemented as a computer program on a Lisp Machine at the MIT Artificial Intelligence Laboratory. The computation made use of a relatively efficient variable–metric minimization procedure developed by Davidon (1968). For a quadratic function of n variables, this method is guaranteed to converge to a minimum within at most n iterations. The computational load at each stage is relatively small, estimated by Davidon (1968) to require approximately $\frac{3n^2}{2}$ multiplications. When the objective function (in our case, the overall deviation from rigidity) has more than a single minimum, the minimization process will converge to a local, but not necessarily the global, minimum. The results described in section 3 demonstrate that this convergence is sufficient for the recovery of the unknown 3–D structure. Some consequences of the convergence to the local minimum are discussed in section 4.4.

For the flat initial model, an additional step is required to ensure convergence to a local minimum. The flat internal model can change into two equally likely configurations, one being the mirror image reflection of the other about the image plane. The model is therefore perturbed slightly, to cause it to prefer one of the two symmetric minima over the other.

This minimization method is efficient for implementation on a serial digital computer. More parallel distributed implementations are also possible. Such extensions will not be analyzed here mathematically (for a discussion of minimization in distributed networks see Ullman 1979b). Instead, a mechanical analogue that performs essentially the same computation in a parallel distributed manner will be briefly described. This mechanical spring–model (which bears some similarity to Julesz' spring-dipole model of stereopsis (Julesz, 1971)) can help to visualize the computation performed at each step by the incremental rigidity scheme, and can be helpful in suggesting parallel distributed computations for maximizing rigidity.

The mechanical spring model is illustrated in figure*2 for a three–element object. As before, let (x_i, z_i) denote the image coordinates of the i^{th} point, and y_i be the unknown depth coordinates that must be recovered. This situation is modeled in figure 2 by a set of rigid rods, each one connected to one of the viewed points, and extending perpendicular to the image plane. The i^{th} point is constrained to lie along the i^{th} rod, but its position along this rod (i.e., its depth value) is still undetermined. The points are now connected by a set of springs. The resting length of the spring connecting points i and j is L_{ij}, their distance in the internal model prior to the introduction of the new frame, and k_{ij} is the spring constant. The points will now slide along the rods, stretching some of the springs and compressing others, until a minimum energy configuration is reached. If l_{ij} denotes the distance between points (i, j) in the final configuration, then the total energy of the system would be $\frac{1}{2}\sum_{i,j} k_{ij}(L_{ij} - l_{ij})^2$.

To mimic the computation described in the preceding section, the spring constants k_{ij} should be smaller for longer springs (it can also be assumed that each point is connected only to a number of its nearest neighbors).

The "computation" of the most rigid transformation is performed in this mechanical system in a parallel distributed manner. It can be used, therefore, to illustrate the possibility of maximizing rigidity in the observed transformation using a parallel network of simple interacting computing elements.

This mechanical analogue illustrates the computation for the case of orthographic projection. For perspective projection, only a slight modification is required: the rods should converge to a common point rather than be perpendicular to the image plane. Continuous versions of this scheme, in which the rods move continuously and the springs' lengths and constants are also modified continuously are possible, but they will not be discussed further here.

3. Recovery of the 3–D structure by the incremental rigidity scheme

3.1 Rigid motion

Typical results showing the incremental rigidity scheme in operation are illustrated in figure 3. The object in this example is shown in figure 3a. It contains six points: the vertices of the outlined pentagon, and a sixth point at the origin (marked by the unfilled circle). The object is shown from a top view, i.e. as projected on the $X - Y$ plane. The input to the incremental rigidity program consisted of the projection of the object on the $X - Z$ image plane. That is, only the (x_i, z_i) coordinates for $i = 0, \ldots, 5$ were given. This projection on the $X - Z$ image plane

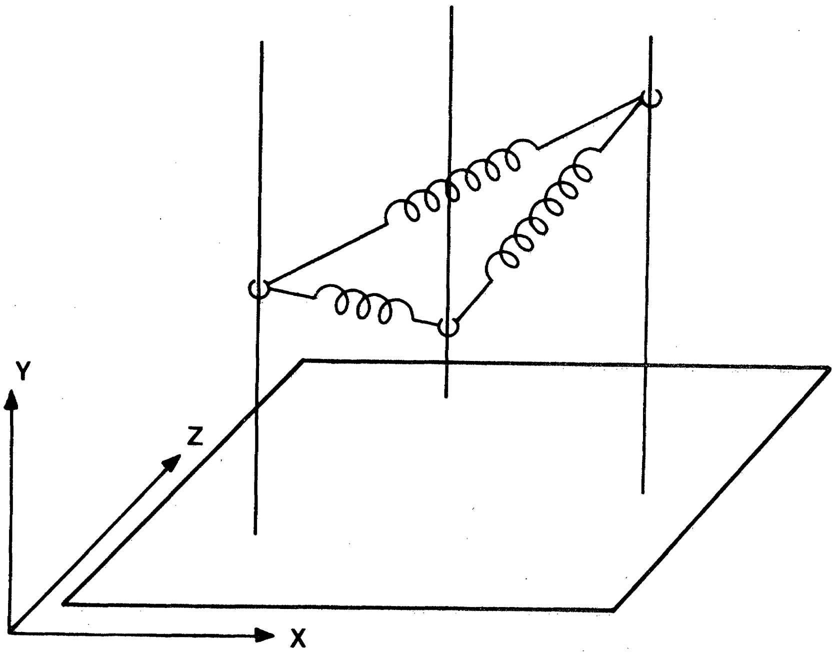

Figure 2 A spring model for the distributed computation of the most rigid interpretation. Each of the viewed points (three in this example) is constrained to move along one of the rigid rods along, and its position along the rod represents its depth value. The connecting springs represent the distances between points in the current internal model. The points would slide along the rods until a minimum energy configuration is reached. The final configuration represents the modified internal model.

at time $t = 0$ is shown in figure 4. The unknown depth values y_i were assumed initially to be constant, $y_i = 0$ for $i = 0, \ldots, 5$. That is, no depth was assumed, and the initial internal model consisted of a planar object, lying parallel to the image plane. (The dashed line in 3a illustrates the projection of the internal model onto the $X - Y$ plane.)

The object was then rotated by 10^0 at a time, and the internal model was modified according to the scheme described in section 2.1. The rotations were around the vertical Z axis. Any other axis in space can be used instead, however, and the axis may also change over time. To illustrate the behavior of the scheme, the error between the internal model and the object's correct 3–D structure was computed at the end of each step. This error was measured as $(\sum_{i,j}(d_{ij} - L_{ij})^2)^{\frac{1}{2}}$ where d_{ij} is the correct 3–D distance between points i and j in the object, and L_{ij} is the corresponding distance in the internal model. When the internal model is completely accurate, $L_{ij} = d_{ij}$ for all i, j and the total error vanishes. The initial error was normalized to 1.0 at 0^0 rotation. Its development as a function of the rotation angle is shown in figure 5 (dotted curve). It declines to about 0.3 after the

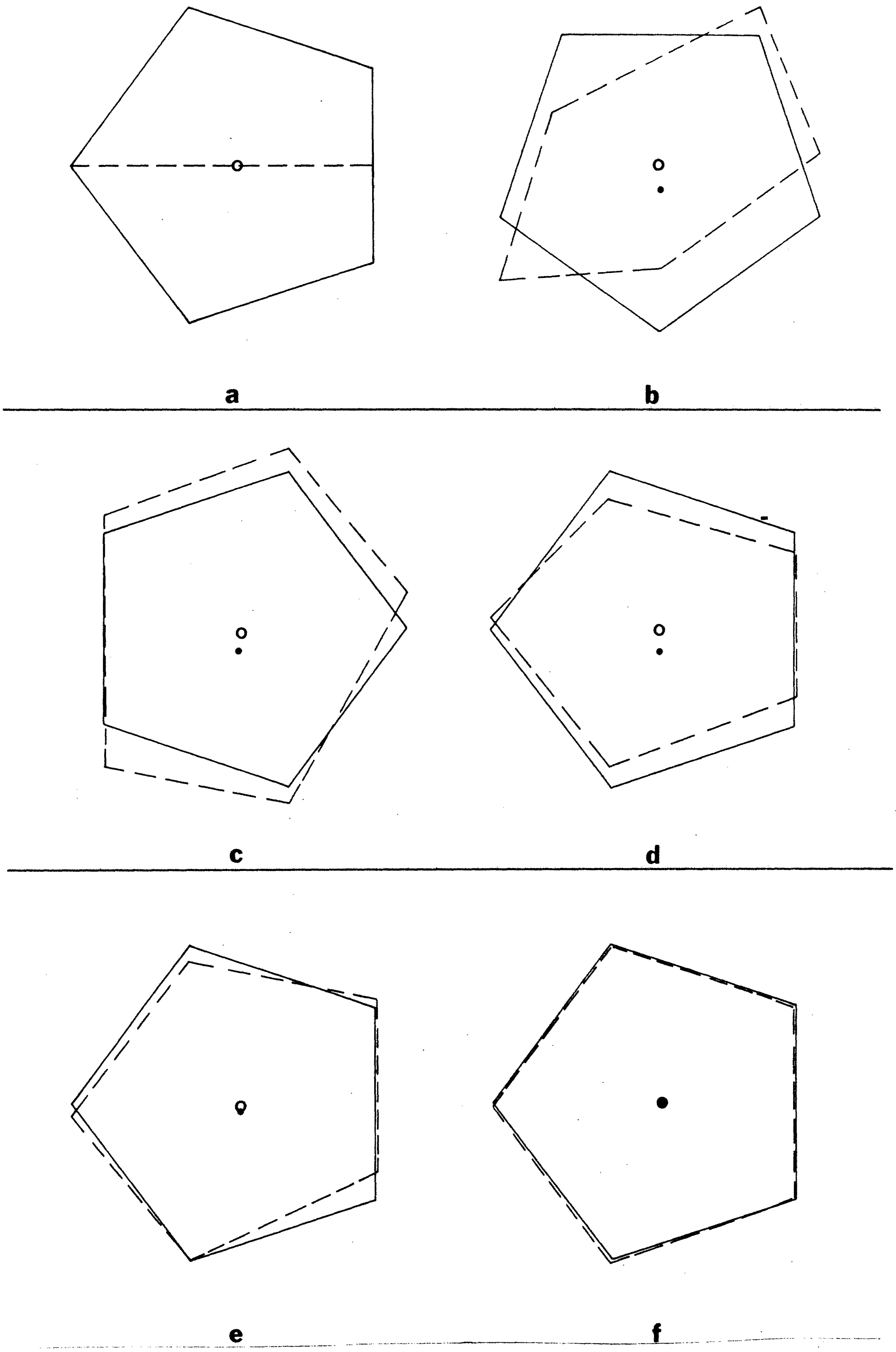

Figure 3 The incremental rigidity scheme applied to a six-point rigid object (the solid outlined pentagon and unfilled dot at the center). The internal model (dashed curves and filled dot near the center) is compared to the correct structure following (a) 0 rotation, (b) 90^0, (c) 180^0, (d) 360^0, (e) 2 rotations, (f) 4 rotations.

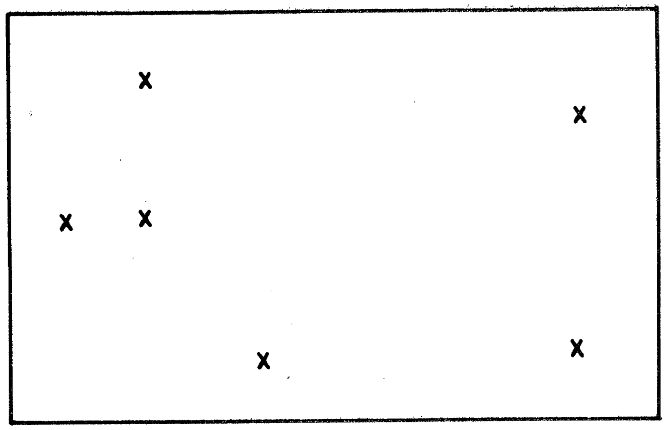

Figure 4 The initial projection of the 6-point object on the image (x-z) plane.

first $180^0$of rotation, (note in fig. 3c that this error already yields an approximation to the actual structure) and then continues to decline to about $2 - 5\%$. Figure 3 shows the development of the internal model. It starts as entirely flat (3a). After 90^0 of rotation it acquires some depth, but it is still too flat, and the shape is inaccurate (3b). After the first 180^0 of rotation the internal model is already similar in overall shape to the correct structure (3c). The internal model continues to improve (fig. 3d,e) until it becomes virtually indistinguishable from the correct 3-D structure (3f).

A more rapid approximation to the correct 3-D structure can be obtained by using the flexible model scheme described in section 2.2. The continuous curve in figure 5 illustrates the error measure as a function of rotation angle for the flexible model scheme. It can be seen that the approximation improves rapidly over the first 180^0 of rotation, but it remains somewhat more oscillatory than the basic scheme.

The results of applying the incremental rigidity scheme to various objects in motion show that for most of the rotation time the internal model approximates the actual 3-D structure. The model does not converge, however, to the precise solution, but often wobbles somewhat around the correct solution to the 3-D structure. In both the basic and the flexible model schemes the approximation to the correct solution does not improve monotonically as a function of rotation angle. The residual non–rigid deformations often increase and then decrease again. The lack of monotonicity in the overall convergence to the computed 3–D structure suggests that an analytic mathematical treatment of the convergence properties of the incremental rigidity scheme is probably difficult to obtain.

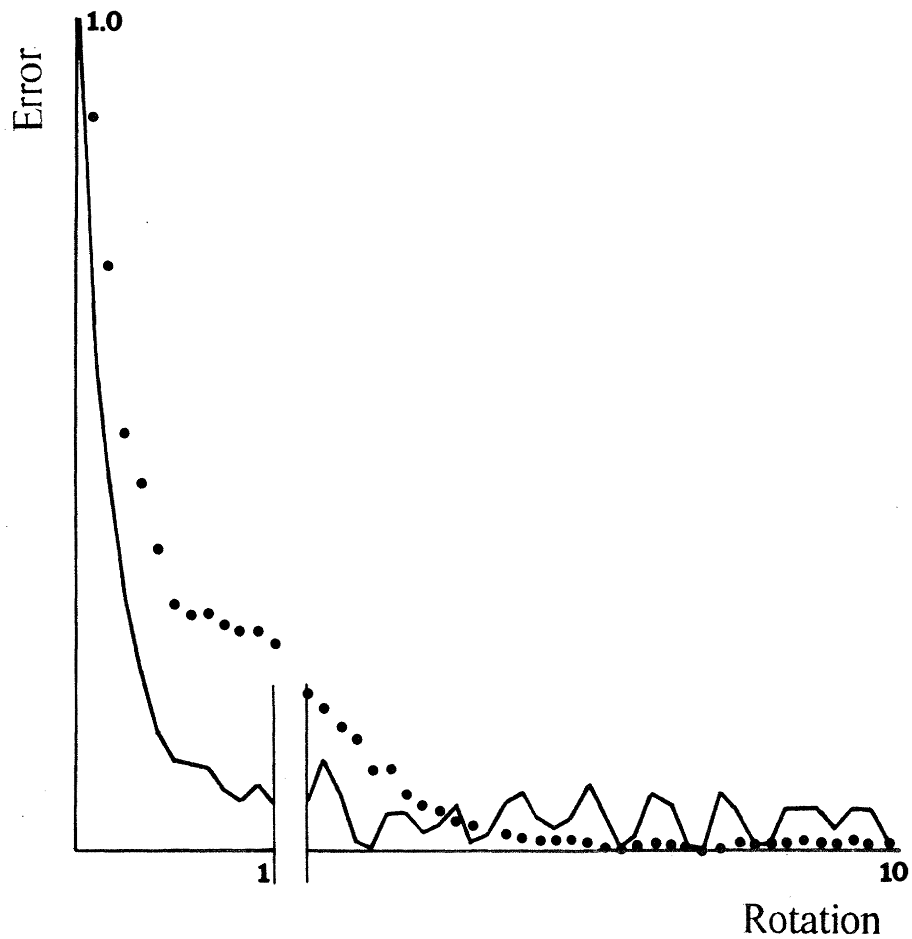

Figure 5 The error of the internal model as a function of rotation angle over 10 revolutions. The initial error for the basic scheme (dots) and flexible model (solid curve) is normalized to 1. The error is shown every 30^0 for the first revolution, and every 90^0 afterwards.

During these oscillations of the error function, the correct structure is occasionally lost temporarily and then recovered. In the course of such a phase, when the structure is lost and recaptured, a spontaneous depth reversal may occur. That is, the internal model converges not to the original 3-D structure, but to its mirror image, reflected about the image plane. The convergence to the reflected rather than to the correct structure is a "legal" solution under orthographic projection, since, as noted in section 1.2, the two are indistinguishable under orthographic projection. (This point of depth reversals under orthographic and perspective projections is discussed further in section 4.1).

Figure 6 shows the development of the internal model for a five point object, similar to the six point object examined before, but without the point at the origin. Figures 6a,b, and c compare the internal model (dashed line) with the actual structure (solid lines) following 0, 90, and 180^0 respectively. Towards the end of the second revolution the structure was temporarily lost, and then recovered successfully. During this phase, a depth reversal occurred. That is, the internal model later converged not to the correct 3–D structure, but to its mirror image, reflected about the image plane. The approximation to the reflected structure eventually becomes quite accurate. Figure 6d shows the correct structure together with the best approximation obtained within the first five revolutions. Figure 6e is similar to 6d, but the correct structure has been reflected about the image plane. It can be seen that the internal model provides a good approximation to the reflected structure. The best approximation obtained within the first ten revolutions is compared in figure 6f against the correct (but reflected) structure. That the structure is recaptured following a total loss, together with the initial convergence from a totally flat internal model, indicates that almost irrespective of the initial conditions the scheme eventually converges, in the sense that it spends most of its time near the correct solution.

In the examples above the objects have been rotated 10^0 between successive views. It might have been expected that if a sequence of frames is taken, say, every 5^0 of rotation instead of the 10 used above, the recovery of the structure would require a smaller overall rotation, since the deviation from rigidity at each step is smaller. In fact, when smaller angular separations between views are used, the convergence becomes somewhat slower. Figure 7 compares the decline of the error function over the first five rotations for 10^0 (dotted curve) and 5^0 (continuous curve) rotations between successive views. This difference in convergence rate suggests that the incremental rigidity scheme performs better when successive views of the object differ significantly. This preference may be related to the findings of Petersik (1980) who compared the contribution of the short– and long–range motion processes (Braddick 1974) to the recovery of structure from motion. The long–range process, which operates over relatively large spatial and temporal separations, was found in this study to be the main contributor to the structure from motion process.

Finally, comparisons were made with the type of to–and–fro motion used in the original kinetic depth experiments by Wallach & O'Connell. In the examples examined above, the objects rotated continuously in one direction for several rotations. In contrast, the objects in Wallach & O'Connell's experiments were rotated to–and–fro through a limited angular excursion. Under this condition, the observers did not have the benefit of viewing the object from all directions, but they were nevertheless able to recover the correct 3-D structure of the moving objects. A simulation of this condition, in which the objects were rotated by only 40^0 in each direction, revealed that the incremental rigidity scheme manifests a similar capacity. As the object rotated to–and–fro, the internal model continued to improve until the correct 3-D structure was recovered, in a manner analogous to the recovery of the 3-D structure of continuously rotating objects.

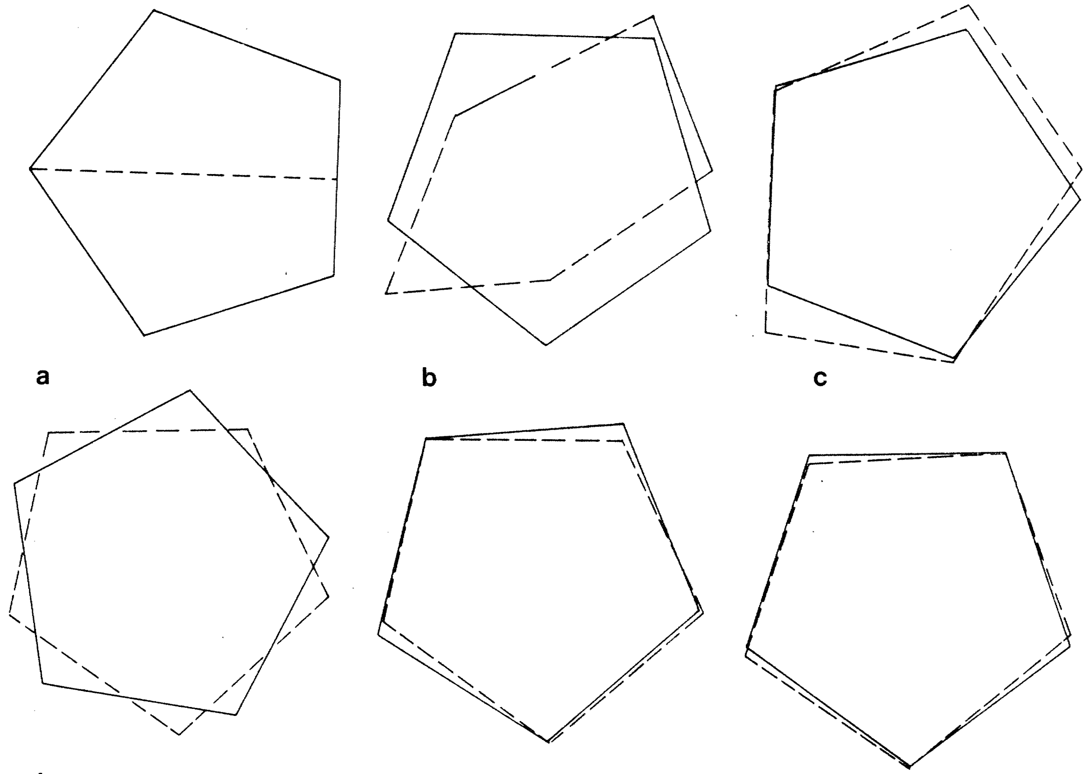

Figure 6 The internal model (dashed figures) is compared to the correct structure (solid pentagons) following (a) 0^0, (b) 90^0, (c) 180^0, (d-e) 5 revolutions, (f) 10 revolutions. Towards the end of the second revolution, the structure was temporarily lost. During this phase, a depth reversal occurred (d). Figures (e) and (f) compare the internal model to the correct 3-D structure reflected about the image plane.

In summary, when applied to rigid objects in motion, the incremental rigidity scheme exhibits the following properties. (1) Veridicality: for most of the time a reasonable approximation of the correct 3–D structure is maintained. (2) Temporal extension: the time (number of frames) required for an approximation to be obtained is longer than the theoretical minimum required for the recovery of structure from motion (Ullman 1979, Longuet–Higgins & Pradzny 1980, Tsai & Huang 1982). (3) Residual non–rigidity: although the changing image is induced in this case by completely rigid objects in motion, the computed 3–D structure included residual non–rigidity. (4) Non–monotonicity: Starting from a flat internal model, the solution generally improves with time. This improvement is, however, non–monotonic. The error often increases and then decreases again. (5) Depth reversals: occasionally the increased error is associated with a spontaneous depth reversal. The flexible model scheme is less susceptible to such reversals.

Similar general properties are also manifested in the perception of structure

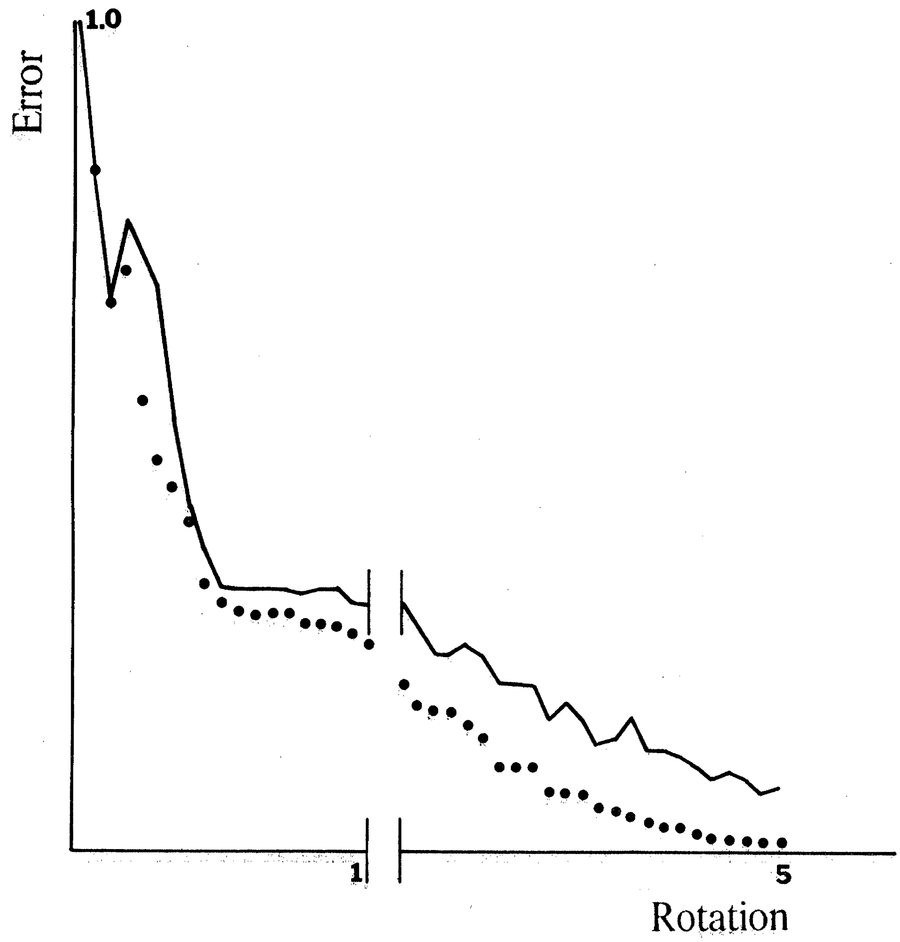

Figure 7 The decline of the error function for a rotation rigid object when successive frames are separated by 5^0 (solid) and 10^0 (dashed curve) of rotation. The convergence is faster when successive views of the object differ significantly.

from motion by human observers. The perceived 3–D structure is usually similar to the correct 3–D structure. It improves with time, but it is usually not entirely accurate (Wallach & O'Connell 1953, White & Mueser 1960). The perception is often of a stable 3–D configuration accompanied by some residual elastic deformations, particularly when the number of participating elements is small.

3.2 Non–rigid motion

In this section the capacity of the incremental rigidity scheme to cope with deviations from rigidity will be examined. Unlike the previous section where the viewed objects were assumed to be entirely rigid, in this section they are allowed to deform while they move.

An example of the scheme applied to non–rigid motion is shown in figure 8. At time $t = 0$ the object was identical in shape to the five point object examined under rigid motion in the previous section. A non–rigid transformation was now added to the rotation of the object. The non–rigid distortion was quite significant, as can be seen in figure 8a. The shape of the object following two revolutions is compared in the figure with its original shape. The incremental rigidity scheme copes successfully with such deviation from rigidity. The internal model by the end of the second revolution is shown in fig. 8b and compared with the correct structure.

Figure 9 illustrates the results of applying the same scheme to an object distorting twice as fast. That is, the distortion of the object after one revolution is identical to the distortion spread in the previous example over two revolutions. The figure compares the actual object with its internal model after 180 (a), 360 (b), 450 (c, where the error is relatively low) and 720^0 of rotation (d, where the error is relatively high again). The internal model becomes less accurate compared to the lower distortion rate, but the 3–D structure is still essentially recovered.

When the rate of distortion was doubled again, the incremental rigidity scheme could no longer cope with the rate of deviation from rigidity. Before the end of the second revolution the structure was lost entirely. The distortion was evidently developing too fast to allow the scheme to recover from this loss. This limitation held for both the basic and the flexible model schemes. In contrast with previous cases, the error measure in this case tends to grow without bounds.

Different objects under different distortions were also examined, with similar results. For moderate distortions the incremental rigidity scheme can cope successfully with non–rigid motion. The amount of distortion that can be tolerated is difficult to quantify, but as illustrated in figures 8 and 9 it can be substantial.

As noted in section 1.3, the human visual system can also cope to some degree with kinetic depth effects that are not entirely rigid. Although no systematic studies of this capacity have been reported, the human visual system is probably susceptible to similar difficulties with non–rigid motion. That is, it is expected to fail under pure kinetic depth conditions (i.e., when no other sources of 3–D information are available) when the deviation from rigidity becomes excessive (c.f. Lappin *et al* 1980). It may be of interest to investigate further the performance of the human visual system under non–rigid conditions and compare the results with the performance of the incremental rigidity scheme.

4. Additional properties of the incremental rigidity scheme

This section discusses four additional topics pertaining to the incremental

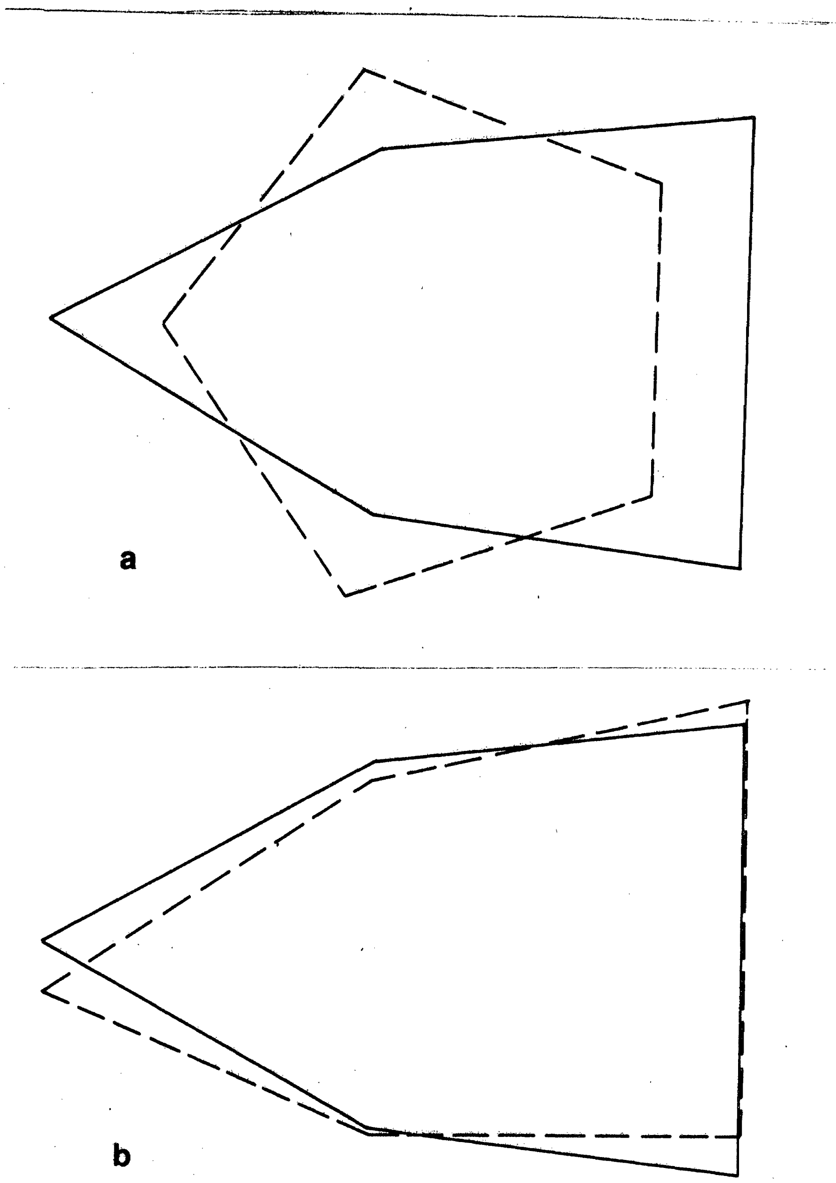

Figure 8 The recovery of non-rigid shape. A pentagon distorts while it moves. (a) Its shape following 2 revolutions (solid) is compared with its original shape (dashed lines). (b) The internal model by the end of the second revolution (dashed) is compared with the correct 3-D structure seen from a top view (solid lines).

rigidity scheme under both rigid and non–rigid motion. Within each topic, the scheme is compared to previous mathematical models and to human perception of structure from motion.

4.1 Orthographic and perspective projections

The computations described above used orthographic projection. As noted in section

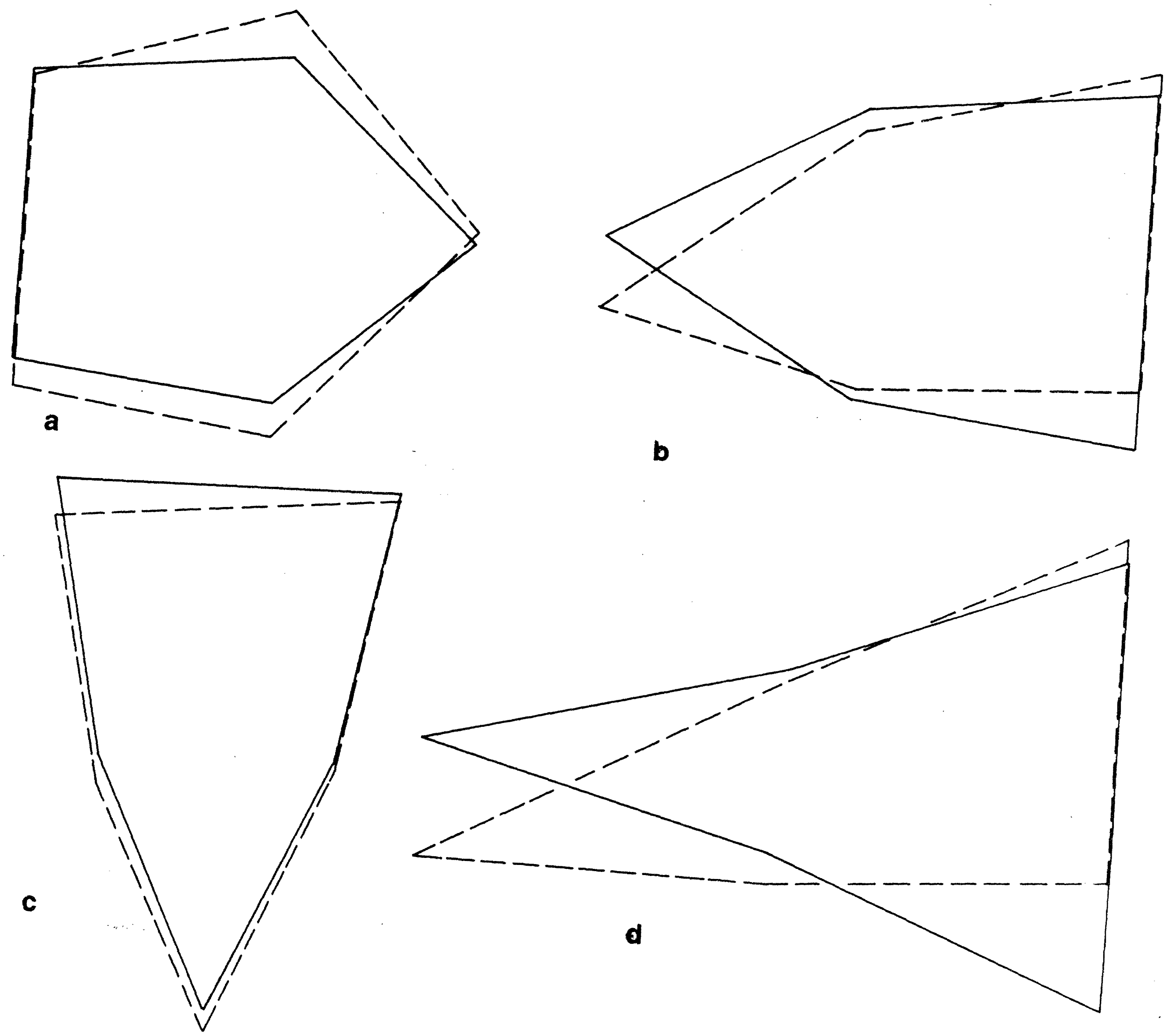

Figure 9 An object distorting twice as fast as that in Figure 8. The internal model (dashed line) is compared to the correct structure following (a) 180^0, (b) 360^0, (c) 450^0, (d) 720^0 of rotation. When the rate of distortion increases further, its structure can no longer be recovered by the incremental rigidity scheme.

2.3, a similar scheme can also be adapted for perspective projection. For small objects (compared to the overall viewing distance) the two projections are similar. If the interpretation is applied locally, it is not crucial which projection is assumed, since the perspective effects would be too small to be used reliably. For the human visual system, Johansson (1978) reports that for objects subtending up to about 15^0 of visual angle perspectivity has a very limited effect. Under such conditions

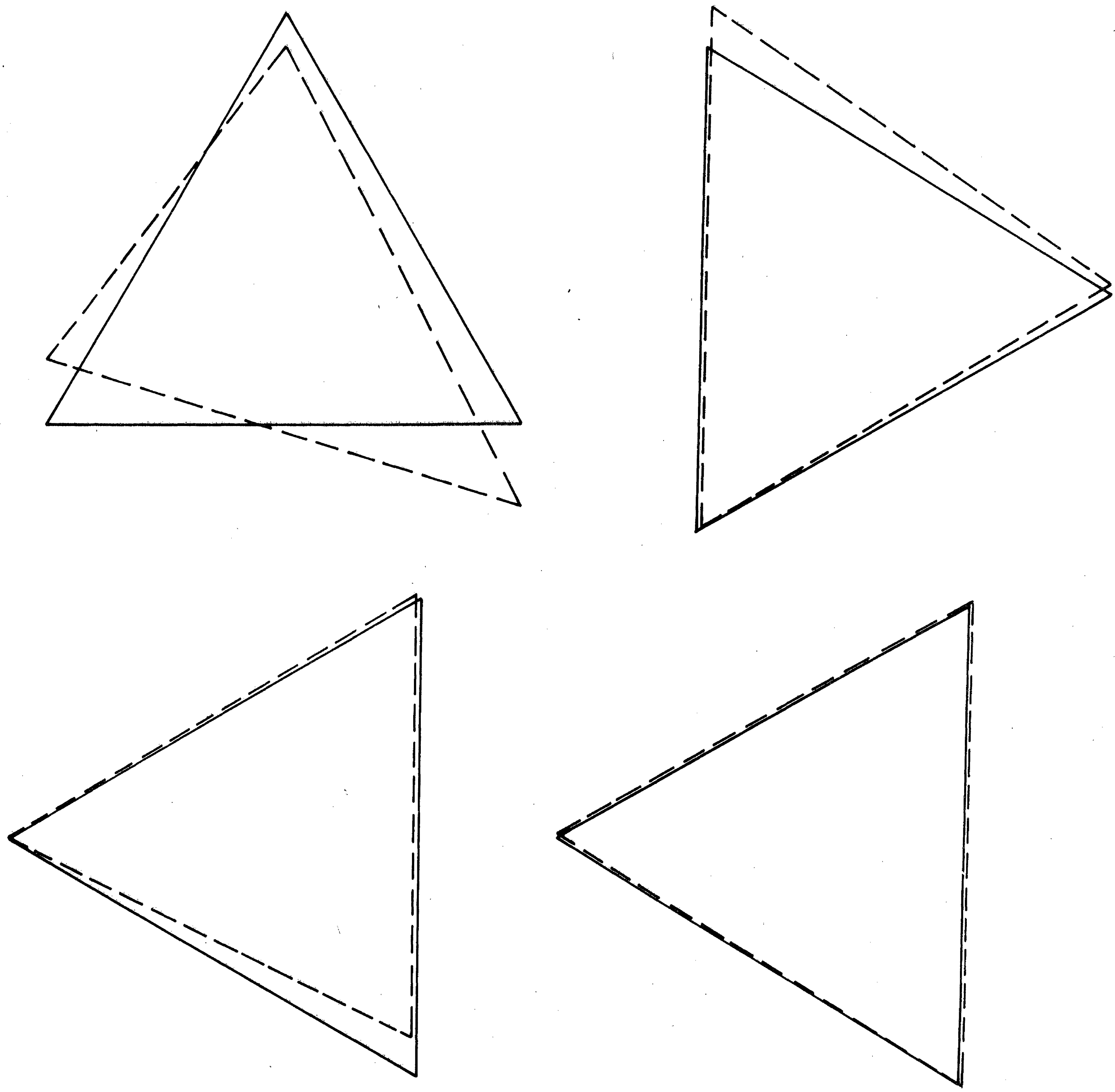

Figure 10 The incremental rigidity scheme applied to a 3-point object. The internal model (dashed lines) is compared to the correct structure (solid lines) as seen from a top view, following (a) 90^0, (b) 180^0, (c) 360^0 and (d) 720^0 of motion. Three points are sufficient for the recovery of structure from motion.

orthographic projection can be viewed as slightly distorted perspective projection, and since the recovery scheme should be insensitive to small distortions, it should be able to cope with both types of projection. In fact, the capacity to deal successfully with both types of projection can be used as a test for the scheme's robustness. A scheme that can recover the structure under perspective projection but fails under orthographic projection cannot be robust when applied locally. This comment is relevant in particular to schemes that rely exclusively on the instantaneous velocity field, since such schemes fail under orthographic projection.

For larger objects, perspective and orthographic projections differ. It is still possible, however, to use a parallel scheme (Ullman 1979) in which the interpretation is performed locally (and therefore it is immaterial whether orthographic of perspective projections are employed), and the local results are then combined in a second stage. For sufficiently large objects this integration stage will eliminate the ambiguity inherent in orthographic projection regarding direction of rotation and reflection about the image plane. In summary, for small objects or surface patches it is immaterial which of the two projection types is employed, and a robust recovery method should be able to cope with both. For larger objects it is still theoretically possible to use either type, and either one can be incorporated in the incremental rigidity scheme.

4.2 The effect of number of points

In this section the effect of the number of moving feature points will be discussed by comparing its application to two, three, four, and many points in motion.

Two points in motion: For two points the 3–D structure is not determined uniquely by any number of views. The structure imposed by the incremental rigidity scheme would be of a rigid rod rotating in depth, and the view where the rod's length is maximal would be taken as lying in the frontal plane. This 3–D interpretation is in agreement with human perception of two–dot configurations (Johansson & Jansson 1968).

Three points: This configuration has not been analyzed mathematically in the past. It is known that four points in three views determine the 3–D structure uniquely if the structure is assumed to be rigid. Three points in three views do not always guarantee uniqueness, but it is still possible that with additional views the 3–D structure is determined uniquely by three points alone. The results of applying the incremental rigidity scheme support this possibility, since the 3–D structure of three–point configurations can be successfully recovered.

The recovery of the 3–D structure of three moving points is shown in figure 10a. As before, the initial model was taken as entirely flat. The evolving internal model (dashed line) is compared in the figure with the actual 3–D structure following 90, 180, 360 and 720^0 of rotation. The figure shows that a fast and accurate recovery can be obtained for only three points in motion.

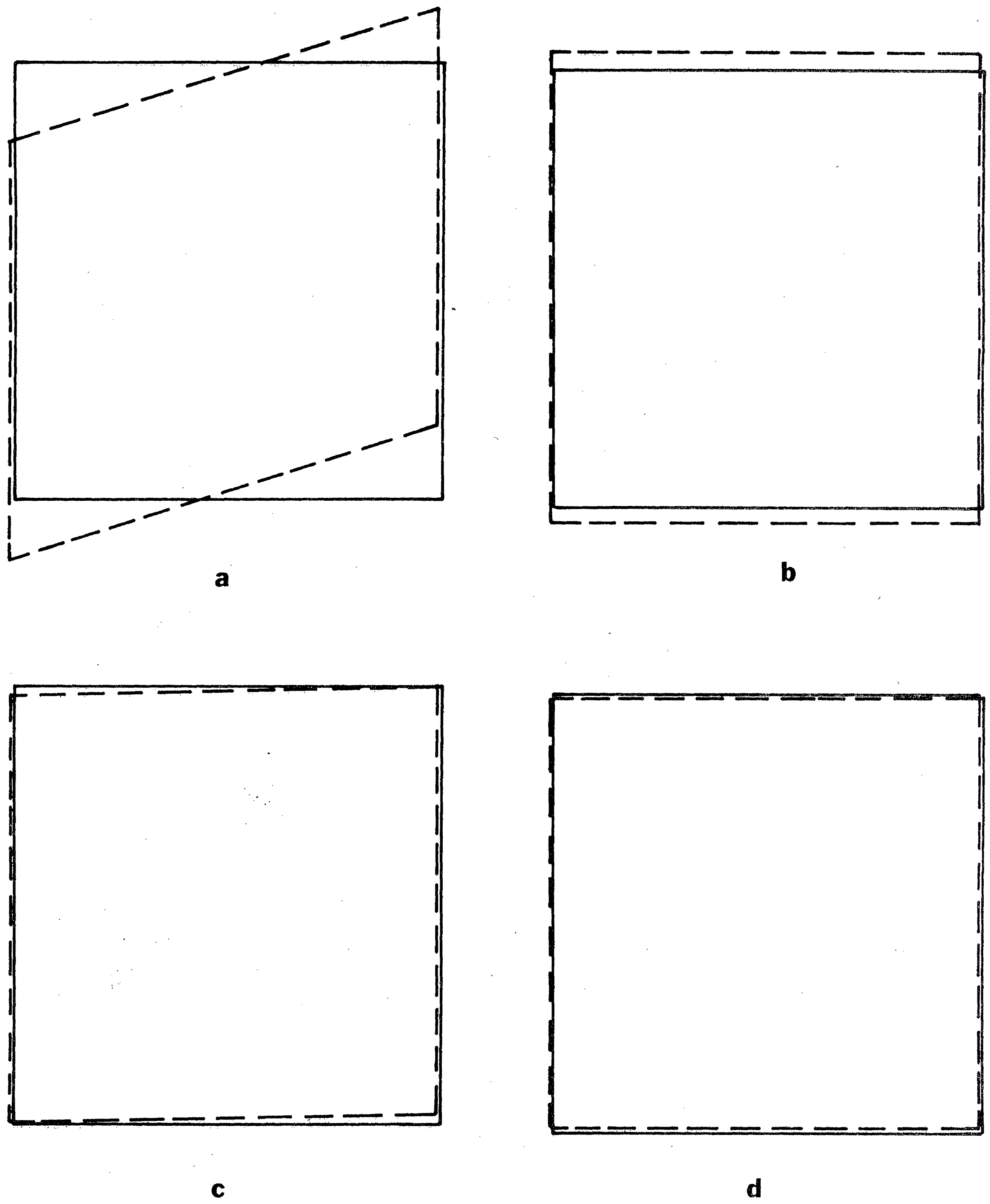

Figure 11 The recovery of the 3-D structure of 4 points arranged in a square when seen from a top view. The initial model was entirely flat. The model (dashed) is compared to the actual structure following (a) 90^0, (b) 180 ^{0}, (c) 360^0, and (d) 720^0 of rotation. The recovery takes longer than the known minimum of 3 views, but the correct structure is eventually recovered and maintained.

Four points: Three views are theoretically sufficient for recovering the structure of four (non–coplanar) points (assuming rigidity). The incremental rigidity scheme can recover the structure of four–point objects, but three views are insufficient for an accurate recovery.

The recovery of the 3–D structure of four points arranged in a square when viewed from a top view is shown in figure 11. The initial model was taken again as entirely flat. The internal model (dashed line) following 90, 180, 360 and 720^0 of rotation is compared to the actual 3–D structure. It can be seen that the model is initially inaccurate, but that the correct 3–D structure is eventually recovered and retained with only minor residual deviations from rigidity.

Additional points: The effect of additional points on the incremental rigidity scheme would depend on the method of applying the scheme to large collections of elements. There are two possible methods of applying the scheme to such large collections. The first is a single–stage scheme, in which the computation described in section 2 is simply applied simultaneously to all of the elements in view. The second possibility is a two–stage scheme (similar to the polar–parallel scheme described in Ullman 1979a). In the first stage the incremental rigidity scheme is applied independently to small subcollections of elements. In the second stage the local results are combined. It is expected that for the two–stage scheme there would be a more noticeable improvement with the number of elements (c.f. Braunstein 1962, Johansson 1978).

The effects of numerosity would also depend on the function D, used in measuring the deviation from rigidity (section 2). If it falls off rapidly as a function of spatial distance, only the nearest neighbors would make substantial contribution to the computation, and the effect of numerosity would be more restricted compared to a function that falls off more gradually with distance.

4.3 On multiple objects

One possible method of dealing with multiple independently moving objects, is to segregate the scene into objects (e.g. on the basis of distance, 2–D common motion characteristics, etc.) before applying the rigidity–based interpretation scheme to each object separately. An interesting alternative is that object segregation may not be required as a separate stage, but may be a by–product of the interpretation process. Similar to the previous section, two methods for achieving such segregation seem possible. First, the segregation may be accomplished in a single stage process by the appropriate choice of the deviation measure D. The suggestion is that D would prefer "partial rigidity" in the following sense. Suppose that no rigid transformation of the internal model can account for the incoming input. The model must then be modified non–rigidly. For simplicity, assume that only two different modifications of the model are possible. In the first all the internal distances are changed somewhat. The second maintains partial rigidity: some distances change more than in the first deformation, but others remain completely rigid. We would want the measure of deviation from rigidity to be lower for this second, partially rigid, transformation. For two independently moving objects, a scheme that maximizes rigidity would then prefer the solution in which the scene contains two rigid substructures. In this manner, the appropriate choice of D may endow the incremental rigidity scheme with the capacity to divide the scene into its rigid components.

A second possibility for dealing with multiple objects is within the framework of the two–stage process mentioned above. The general suggestion is that substructures that share similar motion parameters (e.g., same axis of rotation) in the first stage, will be grouped together in the second stage. This method is similar to the one described in (Ullman 1979b) for completely rigid objects, in that it places the burden of the segregation process on the integration stage.

4.4 Convergence to the local minimum

The schemes described in sections 2.2 and 2.3 seek the local minimum in the measure of deviation from rigidity. As illustrated by the examples examined in section 3, this convergence to the local minimum is usually sufficient to recover the correct 3–D structure. Under certain conditions, however, the incremental rigidity scheme may converge to a local minimum which is not the most rigid structure possible. Similar behavior is also exhibited by the human perception of structure from motion. Under certain conditions human observers perceive non–rigid structure in motion when an entirely rigid solution is also possible. A well known example of this phenomenon is the Mach illusion. (Mach originally described a static version of this illusion. The dynamic version is described in Eden 1962, Lindsay & Norman 1972.) This illusion can be created by folding a sheet of paper to create a vertical V–shaped figure. Under monocular viewing, this shape is ambiguous and can reverse in depth. To observe the dynamic illusion the observer waits for a depth reversal to occur, and then moves his head in different directions. Under these conditions the object is seen to move whenever the observer's head moves. The illusory motion arises despite the observer's knowledge of the correct 3–D configuration, and it often contradicts shading clues, stability criteria, and touch cues (Eden 1962). When the object is close to the observer's eye, its motion is no longer rigid, but appears to distort considerably while it moves.

The incremental rigidity scheme would also be susceptible to this illusion. The reason lies in the initial internal model. Unlike the pure kinetic depth situation, a 3–D structure is perceived from the static view. Because of the depth reversal, the initial internal model resembles the reflection of the 3–D structure about the image plane. It subsequently converges not to the correct 3–D structure, but to its mirror image which, under perspective projection, can be considerably less rigid than the correct structure.

Sperling *et al* (1983) have shown that when one face of a rotating wire cube is increased in brightness, the cube is often perceived as non-rigid. This behavior is consistent with the incremental rigidity scheme. The brighter face is usually perceived as closer to the observer even when in fact it may be the farther. This bias of the internal model will cause the incremental rigidity scheme to converge to the reflected structure and miss the correct, entirely rigid, solution.

5. Summary

A new scheme was suggested for the recovery of 3–D structure from rigid and non–rigid motion. According to this scheme an internal model representing the 3–D structure of the viewed object is maintained and modified as the object moves with respect to the viewer, or changes its structure. The transformations in the internal model thus mirror the changes in the environment, similar to the manner suggested by Shepard (Shepard & Metzler 1971, Shepard & Cooper 1982). The internal model resists changes in its shape as much as possible. Consequently, of all the modifications of the model that may account for the observed transformation, the most rigid one is preferred.

This method of recovering 3–D structure from motion shares a number of properties with the perception of kinetic depth displays by human observers. The internal model is initially inaccurate and improves with time (successive approximation). In the lack of static 3–D information the model prior to the beginning of the motion may initially be entirely flat. As the object starts to move, the model begins to acquire depth, and eventually it reaches a configuration that approximates the actual 3–D structure (convergence). If the initial view does convey 3–D information, this information may affect the structure of the internal model (integrating sources of information). The recovery process eventually integrates information from different views of the object (temporal extension). The entire history of the process is summarized, however, in the structure of the internal model, the scheme does not operate on long sequences of views or stored trajectories.

The proposed incremental rigidity approach raised two main questions: (1) its convergence to the correct 3–D structure, (2) its capacity to cope with deviations from rigidity. The computer simulations have demonstrated that the use of the instantaneous model alone, coupled with a principle of maximizing rigidity, is sufficient for the recovery of 3–D structure from motion. The resulting scheme has an inherent preference for rigid transformations, but it can also cope with considerable deviations from rigidity.

The simulations also revealed the main advantages and disadvantages of the incremental rigidity scheme. Compared with previous approaches, one limitation of the scheme is that the resulting 3–D structure is usually not entirely accurate (although it often approximates the correct structure quite closely). Even for strictly rigid objects, the computed 3–D solution usually contains residual non–rigid distortions. On the other hand, two advantages of the incremental rigidity scheme make it an attractive approach to the recovery of structure from motion. The first is its capacity to cope with non–rigid motion: the 3–D structure can be approximated in the face of substantial deviations from rigidity. The second is its robustness: errors in the measured velocity and in the computations employed do not result in complete failure, but in some additional non–rigid distortions superimposed on the correct 3–D structure.

A number of problems remain open for further studies. Mathematically, it would be of interest to analyze the convergence properties of the incremental rigidity scheme. As noted in section 3.1, one complicating factor is the lack of monotonicity in the decline of the residual error.

The error function has been defined as a root mean square measure, using the differences in the internal distances in the correct 3–D structure and the internal model. It remains possible that a different error measure would prove more amenable to analytic treatment.

From a psychological point of view, there is a qualitative similarity between the perception of structure from motion by humans, and the behavior of the incremental rigidity scheme. Quantitative data regarding the perception of structure from motion under deviations from rigidity are, however, scant. It would be of interest to investigate further this capacity of the human visual system, and compare the empirical results with the behavior of the incremental rigidity scheme.

Acknowledgement: I thank Dr. Ellen Hildreth for her valuable comments and Carol Jean Papineau for her help in preparing this manuscript.

References

Adelson, E.H. & Movshon, J.A. 1982. Phenomenal coherence of moving visual patterns. *Nature,* **300,** 523–525.

Bobick, A. 1983. A hybrid approach to structure-from-motion". *Proceedings of the ACM Interdisciplinary Workshop on Motion,* Toronto, Canada, April, 1983, pp. 91-109.

Braddick, O.J. 1974. A short–range process in apparent motion. *Vision Research,* **14,** 519–527.

Braunstein, M. L. 1962. Depth perception in rotation dot patterns: effects of numerosity and perspective. *Journal of Experimental Psychology,* **64** (4) 415-420.

Braunstein, M.L. 1976. *Depth Perception Through Motion* New York: Academic Press.

Clocksin, W.F. 1980. Perception of surface slant and edge labels from optical flow: a computational approach. *Perception,* **9**(3), 253–269.

Davidon, W.C. 1968. Variance algorithm for minimization. *The Computer Journal,* **10(4),** 406–413.

Eden, M. 1962. A three-dimensional optical illusion. *Quarterly Progress Report,* No. 64, MIT RLE, 267-274.

Fennema, C.L. & Thompson, W. B. 1979. Velocity determination in scenes containing several moving objects. *Computer Graphics & Image Processing,* **9,** 301–315.

Gibson, J.J. & Gibson, E.J. 1957. Continuous perspective transformations and the perception of rigid motion. *Journal of Experimental Psychology,* **54** (2), 129–138.

Green, B.F. 1961. Figure coherence in the kinetic depth effect. *J. Exp. Psychol.,* **62(3),** 272–282.

Hay, C. J. 1966. Optical motions and space perception - an extension of Gibson's analysis. *Psychological Review,* **73,** 550-565.

Hoffman, D.D. & Flinchbaugh, B.E. 1982. The interpretation of biological motion. *Biol. Cybern.* **42,** 195–204.

Horn, B.K.P. & Schunck, B.G. 1981. Determining optical flow. *Artificial Intelligence* ,**17,** 185–203.

Jansson, G. & Johansson, G. 1973. Visual perception of bending motion. *Perception,* **2,** 321–326.

Johansson, G. 1964. Perception of motion and changing form. *Scan. J. Psychol.,* **5,** 181–208.

Johansson. G. 1973. Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics,* **14** (2), 201-211.

Johansson, G. 1974. Visual perception of rotary motion as transformation of conic sections – a contribution to the theory of visual space perception. *Psychologia,* **17**, 226–237.

Johansson, G. 1975. Visual Motion Perception. *Scientific American,* **232** (6), 76-88.

Johansson, G. 1978. Visual event perception. In: R. Held, H.W. Leibowitz, & H.-L Teuber (ed.), *Handbook of Sensory Physiology* Berlin: Springer–Verlag.

Johansson, G. & Jansson, G. 1968. Perceived rotary motion from changes in a straight line. *Perception & Psychophysics,* **4(3)**, 165–170.

Julesz, B. 1971. **Foundation of Cyclopean Perception.**University of Chicago Press.

Koenderink, J.J. & van Doorn, A.J. 1975. Invariant properties of the motion parallax field due to the motion of rigid bodies relative to the observer. *Optica Acta, 22(9),* 773–791.

Lappin, J.S., Doner, J.F., & Kottas, B.L. 1980. Minimal conditions for the visual detection of structure and motion in three dimensions. *Science, 209,* 717–719.

Lee, D.N. 1976. A theory of visual control of braking based on information about time to collision. *Perception,* **5,** 437–459.

Lee, D.N. 1980. The optic flow field: the foundation of vision. *Phil Trans. Roy. Soc. Lond., B,* **290,** 169–179.

Lindsay, P. H. & Norman, D. A. 1972 **Human Information Processing.** New York and London: Academic Press.

Longuet–Higgins, H.C. 1983. The role of the vertical dimension in stereoscopic vision. *Perception,* **11,** 377–386.

Longuet–Higgins, H.C. & Prazdny, K. 1980. The interpretation of a moving retinal image. *Proc. Roy. Soc. Lond., B,* **208,** 385–397.

Marr, D. & Ullman, S. 1981. Directional selectivity and its use in early visual processing. *Proc. Roy. Soc. Lond. B,* **211** 151–180. Available also as M.I.T. A.I. Memo 524, 1979.

Miles, W. R. 1931. Movement interpretations of the silhouette of a revolving fan. *American Journal of Psychology* **43,** 392-505.

Petersik, J.T. 1980. The effect of spatial and temporal factors on the perception of stroboscopic rotation simulations. *Perception,* **9**,271–283.

Prazdny, K. 1980 Egomotion and relative depth map from optical flow. *Biol. Cyber.,***36,** 87–102.

Shepard, R.N. & Metzler, J. 1971. Mental rotation of three–dimensional objects. *Science,* **171,** 701–703.

Shepard, R. N. & Cooper, L. A. 1982. *Mental Images and Their Transformations.*Cambridge & London: M.I.T Press.

Sperling, G., Pavel, M., Cohen, Y., Landy. M.S. & Schwartz, B.J. 1983. In: O.J. Braddick & A.C. Sleigh (ed.) *Physical and Biological Processing of Images* Berlin: Springer–Verlag.

Tsai, R.Y. & Huang, T.S. 1982. Uniqueness and estimation of three–dimensional motion parameters of rigid objects with curved surfaces. *University of Illinois at Urbana–Champaign Coordinated Science Laboratory Report R-921.*

Ullman, S. 1979. *The Interpretation of Visual Motion* Cambridge & London: MIT Press.

Ullman, S. 1979a. The interpretation of structure from motion. *Proc. Roy. Soc. London, B,* **203**, 405–426.

Ullman, S. 1979b. Relaxation and constrained optimization by local processes. *Comp. Graph. Im. Proc.,* **9(6)**, 115–125.

Ullman, S. 1983. Recent computational results in the interpretation of structure from motion. In: A. Rosenfeld & J. Beck (ed.) *Human and Machine Vision* (in press).

Ullman, S. & Hildreth, E. 1983. The measurement of visual motion. In: O.J. Braddick & A.C. Sleigh (ed.), *Physical and Biological Processing of Images* Berlin: Springer–Verlag.

Wallach, H. & O'Connell, D. N. 1953. The kinetic depth effect. *Journal of Experimental Psychology,* **45**, 205-217.

Wallach, H., O'Connell, D.N. & Neisser, U. 1953. The memory effect of visual perception of 3–D form. *Journal of Experimental Psychology,* **45**, 360–368.

Wallach, H., Weisz, A. & Adams, P. A. 1956. Circles and derived figures in rotation. *American Journal of Psychology,* **69**, 48-59.

Webb, J. A. & Aggarwal, J. K. Visually interpreting the motions of objects in space. *Computer,* **14**, (8), 40-49.

White, B.W. & Mueser, G.E. 1960. Accuracy in reconstructing the arrangement of elements generating kinetic depth displays. *J. Exp. Psychol., 60(1),* 1–11.