MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo 870 **December, 1985**

The Optical Flow of Planar Surfaces

Shimon Ullman

Abstract: The human visual system can recover the 3-D shape of moving objects on the basis of motion information alone. Computational studies of this capacity have considered primarily non-planar rigid objects. With respect to moving planar surfaces, previous studies by Hay (1966), Tsai & Huang (1981), Longuet-Higgins (1984), have shown that the planar velocity field has in general a two-fold ambiguity: there are two different planes engaged in different motions that can induce the same velocity field.

The current analysis extends the analysis of the planar velocity field in four directions: (1) the use of flow parameters of the type suggested by Koenderink & van Doorn (1975), (2) the exclusion of confusable non-planar solutions, (3) a new proof and a new method for computing the 3-D motion and surface orientation (4) a comparison with the information available in orthographic velocity fields, which is important for determining the stability of the 3-D recovery process.

Acknowledgements. This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the Laboratory's Artificial Intelligence research is provided in part by the Advanced Research Projects Agency under Office of Naval Research contract N00014-84-K-0650.

1. Introduction

When objects move in the environment with respect to the viewer, the images they cast upon the viewer's retina undergo complex transformations. The human visual system can interpret these transformations to recover the three-dimensional (3-D) structure of the viewed objects and their motion in space.

A substantial number of computational studies have investigated this capacity to recover structure from motion. The main two goals of these studies have been (i) to determine the conditions under which the 3-D structure can be recovered uniquely from the changing projection, and (ii) to develop methods for computing the 3-D structure and motion in space from the projected transformations.

With respect to the uniqueness problem, the main conclusion from these studies has been that for non-planar rigid objects in motion, the 3-D structure and motion are determined uniquely by the projected transformation. Results of this type have been established for both perspective and orthographic projections, and for different forms of input to the recovery process, including discrete points in discrete views, discrete points and their velocities, and a continuous velocity field.

The uniqueness results obtained in these studies depended criticially upon the object being non-planar. Until recently, the problem of interpreting the motion cast by planar surfaces has remained relatively unexplored. The problem has obvious practical implications, since many surfaces are planar or nearly planar, and in these cases methods that assume non-planarity would be incorrect or unreliable. As mentioned by Waxman & Ullman (1983) and by Longuet-Higgins (1984), the analysis of motion relative to a planar surface may be relevant to certain situations of night-landing of an aircraft, where the main visual cues are roughly coplanar.

Hay (1966) was apparently the first to analyze mathematically the visual interpretation of moving planes. Hay's analysis assumed that the visual input is given in the form of two discrete views obtained from a set of points in motion.

The main result established by Hay was that the interpretation problem in this case exhibits in general a two-fold ambiguity. In addition to the moving plane that has actually cause the viewed transformation there is in general one additional "confusable" plane. The second solution is entirely different from the first in its spatial orientation and motion through space.

Similar results have been established by Tsai & Huang (1981), based on a different method of analysis. They have also shown that when the translattion in space in along the normal to the surface, the ambiguity disappears and the solution becomes unique.

A recent study by Longuet-Higgins (1984) has identified an additional condition under which the two-fold ambiguity disappears. He also provided a different method of analysis and a different algorithm for computing the 3-D parameters.

The current paper extends the analysis of the planar velocity field (i.e., the velocity field induced by a moving planar surface) in four directions. First, it uses a different form of input to the recovery process. Instead of discrete points, it uses continuous flow parameters such as local vorticity, shear, and their derivatives. The use of flow parametsrs in the analysis of the optical velocity field was suggested originally by Koenderink & van Doorn (1975) and by Longuet-Higgins & Prazdny (1980). An analysis based entirely on this form of input has been developed recetly by Waxman & Ullman (1983). When this method was applied to planar surfaces, it was found empirically that there were in general two distinct 3-D solutions. The result was not proven mathematically, however, and the possible existence of additional solutions was not ruled out.

An analysis based on input in the form of flow parameters is of interest for two reasons. First, as emphasized by Koendrink & van Doorn and by Longuet-Higgins& Prazdny, the flow parameters are convenient in the sense that they usually have a clear geometric interpretation, and some of them are invariant with respect to the choice of a coordinate system. Second, in analogy with the non-planar case, it is of interest to examine the interpretation problem for different forms of input since the analysis of each case is usually different, leading to a different recovery method with somewhat different properties.

The second direction in which the current paper extends previous investigations is the consideration of confusable non-planar solutions. Previous studies have assumed that the moving surface is known to be planar, and examined the number of possible planar solutions. It is of interest, however, to determine the ambiguity of the interpretation when the moving surface is not known in advance to be planar. It is shown here that confusable non-planar solutions can in general be ruled out. Third, the method of analysis is different from previous studies, leading to a different family of possible algorithms. Finally, the information available in the orthographic rather than perspective velocity field of planar surfaces is analyzed. Since under local analysis perspective and orthorgraphic projections become almost indistinguishable, this analysis indicates some limits on the information that can be extracted reliably from a local analysis of the velocity field.

2. The Velocity Field of a Moving Surface

The analysis of the planar velocity field will proceed in two steps. In the first step the projected velocity field will be expressed in terms of the 3-D shape and motion parameters of the inducing surface. The descriptoin will follow the derivation in Waxman & Ullman (1983). This step is straightforward, and the resulting 2-D velocity field is obviously uniquely determined by the 3-D parameters. The second step (Sections 3,4,5) consists of inverting this process: given the instantaneous flow field of a planar surface, the problem is to recover the unknown 3-D parameters.

2.1 Notation

Following Longuet-Higgins & Prazdny we will use a coordinate system (X,Y,Z) moving with the observer relative to the scene. The origin of the coordinate system is the vertex of the perspective projection from the scene to the image, and the Z-axis is oriented along the instantaneous line of sight. The Z-axis intercepts the viewed object at (0,0, Z_0). It is assumed that around this point the object can be described by a twice-differentiable surface (not necessarily planar) Z(X,Y). Image coordinates will be denoted by lower-case letters (x, y) where $x = \frac{X}{Z}$, $y = \frac{Y}{Z}$. U, V, W will denote velocities in space in the X,Y,Z directions, and u, v image velocities in the x, y directions. Subscripts such as u_x, u_y will denote partial derivatives along the x and y direction.

2.2 The Instantaneous Velocity Field Around the Origin

The instantaneous motion of a rigid object can be described by six independent parameters. We will denote them by $M_i, i = 1, ...6$. M_1, M_2, M_3 are the velocities along the X, Y, Z directions scaled by the distance $Z_0, M_1 = U/Z_0, M_2 = V/Z_0, M_3 = W/Z_0$. M_4, M_5, M_6 are the angular velocities around the X, Y, Z axes. The shape parameters we will use are $T_1, ..., T_5$, the surface orientation at the origin $T_1 = \left(\frac{\partial Z}{\partial X}\right)_0$, $T_2 = \left(\frac{\partial Z}{\partial Y}\right)_0$, and the surface curvature at the origin (scaled byZ_0) $T_3 = Z_0\left(\frac{\partial^2 Z}{\partial X^2}\right)_0$, $T_4 = Z_0\left(\frac{\partial^2 Z}{\partial Y^2}\right)_0$, $T_5 = Z_0\left(\frac{\partial^2 Z}{\partial x \partial y}\right)_0$. T_i are assumed below to be finite.

The image observables that will be used to recover the unknown 3-D shape and motion parameters are the image velocity at the origin (0,0) and its first and second derivatives in the x and y directions. Rather than using these derivatives themselves, we will use linear combinations of them, as suggested originally by Koenderink & van Doorn (1975) and by Longuet-Higgins & Prazdny (1980).

We will use a set of 12 image observables denoted by $O_1, ..., O_{12}$. They are defined as follows.

$$
\begin{aligned}
O_1 &= u & O_7 &= u_{xx} \\
O_2 &= v & O_8 &= u_{xy} \\
O_3 &= u_x & O_9 &= v_{yx} \\
O_4 &= v_y & O_{10} &= v_{yy} \\
O_5 &= \tfrac{1}{2}(u_y + v_x) & O_{11} &= w_x \\
O_6 &= \tfrac{1}{2}(u_y - v_x) \equiv w & O_{12} &= w_y
\end{aligned}
$$

The first two observables evaluated at (0,0) are the image velocity at the origin in the x and y directions. The next four can be thought of as describing the deformation of a differential neighborhood around the origin. O_3 and O_4 are the rate-of-stretch along the x and y axes, O_5 measures the rate of decrease in the angle between line elements oriented along the axes, and O_6 is the local rate of rotation. $O_7 - O_{12}$ are the spatial derivatives of these variables. By evaluating explicitly the various derivatives it is straightforward to obtain the mathematical relations between the observables $O_1, ..., O_{12}$ and the unknown parameters $M_1, ..., M_6$ $T_1, ... T_5$ (Waxman & Ullman 1983. See also a similar derivation in Longuet-Higgins & Prazdny 1980). The resulting equations are:

$$
\begin{aligned}
O_1 &= -M_1 - M_5 & O_7 &= -2(M_5 + M_3T_1) + M_1T_3 \\
O_2 &= -M_2 + M_4 & O_8 &= M_4 - M_3T_2 + M_1T_5 \\
O_3 &= M_3 + M_1T_1 & O_9 &= -M_5 - M_3T_1 + M_2T_5 \\
O_4 &= M_3 + M_2T_2 & O_{10} &= 2(M_4 - M_3T_2) + M_2T_4 \\
O_5 &= \tfrac{1}{2}(M_2T_1 + M_1T_2) & O_{11} &= \tfrac{1}{2}(-M_4 + M_3T_2 + M_2T_3 - M_1T_5) \\
O_6 &= -M_6 + \tfrac{1}{2}(M_2T_1 - M_1T_2) & O_{12} &= \tfrac{1}{2}(-M_5 - M_3T_1 - M_1T_4 + M_2T_5)
\end{aligned} \tag{1}
$$

These relations can be used for the recovery of the 3-D shape and motion parameters by solving for the M_i and T_i when the O_i are given (measured in the image). Longuet-Higgins & Prazdny (1980) have shown that for non-planar patches a similar set of equations has at most three different solutions. It was found in computer simulations (Waxman & Ullman 1983) that the solution is usually unique, but an analytic uniqueness proof is still lacking. We next turn to the planar case and show that eq. (1) then have in general exactly two distinct solutions.

3. The two-fold ambiguity of planar surfaces

In the planar case the surface curvature parameters T_3, T_4, T_5 in (1) all vanish. The resulting equations are still coupled and non-linear, and the number of distinct solutions is not immediately apparent. In this section it is shown that there are in general two distinct solutions. In addition to the surface that gave rise to the velocity field there is in general one (and only one) additional surface, engaged in a different motion, that can produce an identical velocity field.

3.1 There are in general at least two distinct solutions

When $T_3 = T_4 = T_5 = 0$ the last four equations in (1) are immediately derivable from the preceding eight and can be ignored. If the resulting system of eight equations has a solution, it also has a second solution, that is in general different from the first. This claim is established by giving explicitly a second solution in terms of the first. Let $(M_1, ..., M_6, T_1, T_2)$ be a solution to (1) (with $T_3 = T_4 = T_5 = 0$). A second solution $(\overline{M}_1, ...\overline{M}_6, \overline{T}_1, \overline{T}_2)$ can be derived explicitly as follows:

$$\begin{aligned}
\overline{T}_1 &= -M_1/M_3 \\
\overline{T}_2 &= -M_2/M_3 \\
\overline{M}_1 &= -T_1 M_3 \\
\overline{M}_2 &= -T_2 M_3 \\
\overline{M}_3 &= M_3 \\
\overline{M}_4 &= M_4 - M_2 - M_3 T_2 \\
\overline{M}_5 &= M_5 + M_1 + M_3 T_1 \\
\overline{M}_6 &= M_6 + M_1 T_2 - M_2 T_1
\end{aligned} \tag{2}$$

This solution exists provided that $M_3 \neq 0$. Ths case $M_3 = 0$ is examined in section 4 below. The two solutions are dual in the sense that either one can be used in (2) to obtain the other. In the two solutions the directions of the translation vector and the surface normal are interchanged. If v, n are the translation vector and surface normal in the first solution, then the second $\overline{v}$ points in the direction of n, and $\overline{n}$ in the direction of v.

We next turn to show that there are no more than two distinct solutions. This will be done in two stages. Section 3.2, which is the main step in the proof, shows that if two planar surfaces have identical velocity fields then they have the same value of M_3 (velocity along the line of sight). Section 3.3 establishes that there are at most two solutions that share the same value of M_3.

3.2 Two planes that induce the same velocity field have the same value of M_3

Let π and $\overline{\pi}$ be two planes engaged in motions $(M_1, ..., M_6)$ and $(\overline{M}_1, ..., \overline{M}_6)$ respectively, that induce identical velocity fields (i.e., identical observables in (1)). We can assume that the two planes intersect. Otherwise, $T_1 = \overline{T}_1, T_2 = \overline{T}_2$, which implies either (i)$(M_1, ..., M_6) = (\overline{M}_1, ..., \overline{M}_6)$, or (ii) all of the translation components $M_1, M_2, M_3, \overline{M}_1, \overline{M}_2, \overline{M}_3$, vanish. This latter case, corresponding to pure rotation, is uninteresting since no 3-D information is conveyed by the changing image. Let ℓ be the intersection of π and $\overline{\pi}$. This special line in space participates in two different motions and induces the same (linear) velocity fields. Without loss of generality we can re-orient the coordinate system so that the projection of ℓ coincides with the x axis. The line ℓ can be described now by $Z = ax + Z_0$. The main step in the proof is to consider the intersection line ℓ instead of the two planes. This line has the property that when it participates in motion $(M_1, ..., M_6)$ or $(\overline{M}_1, ..., \overline{M}_6)$, it induces identical velocity fields.

Consider the situation in which the planes no longer exist, only the single line ℓ is moving in space. Let it move with the 3-D motion parameter $(M_1 - \overline{M}_1, ..., M_6 - \overline{M}_6)$. From the original coincidence between the velocity fields of π and $\overline{\pi}$, and since ℓ lies on both planes, it follows that the velocity field projected by ℓ now vanishes, i.e., it satisfies at the origin the equations:

$$
\begin{array}{ll}
v = 0 & u = 0 \\
\frac{\partial v}{\partial x} = 0 & \frac{\partial^2 v}{\partial x^2} = 0 \\
\frac{\partial u}{\partial x} = 0 &
\end{array}
\qquad (3)
$$

We have transformed the problem of the moving planes into a problem concerning a moving straight line. The question is: Under what conditions the velocity field of a moving line, as expressed in eq. (3), vanishes? In the re-oriented coordinate system let us denote

$$
\begin{array}{lll}
M_1 - \overline{M}_1 = V_x & M_2 - \overline{M}_2 = V_y & M_3 - \overline{M}_3 = V_z \\
M_4 - \overline{M}_4 = w_x & M_5 - \overline{M}_5 = w_y & M_6 - \overline{M}_6 = w_z
\end{array}
$$

The five equations in (3) can be expressed in terms of the six motion parameters $(V_x, V_y, V_z, w_y, w_z)$. The derivation is somewhat lengthy, but straightforward. The resulting equations are:

$$
\begin{array}{l}
V_x + Z_0 w_y = 0 \\
V_y - Z_0 w_x = 0 \\
V_z + aV_x = 0 \\
aV_z + Z_0 w_y = 0 \\
aV_y - Z_0 w_z = 0
\end{array}
\qquad (4)
$$

From which it follows that:

$$V_x = 0 \quad V_z = 0 \quad w_y = 0 \quad V_y = Z_0 w_x \quad a w_x = w_z$$

In terms of the original planes π, $\overline{\pi}$, the implication of $V_z = 0$ in that $M_3 = \overline{M}_3$. In the particular coordinates system in which ℓ projects onto the x-axis, it is also true that $M_1 = \overline{M}_1$ and $M_5 = \overline{M}_5$.

3.3 The number of distinct solutions cannot exceed two

Since all the possible planar solutions share the same value of M_3, the proof will be completed by showing that for a given M_3 there are at most two distinct solutions. We proceed along the following plan. If π is a planar solution let ℓ be now the intersection line of π with the frontal plane $Z = Z_0$. We will call ℓ the "tilt line" of π. If we re-orient the coordinate system so that ℓ runs along the x-axis, then in the new coordinate system $T_1 = 0$. From eq. (1), in this reoriented coordinate system $O_3 = M_3$. It can be observed from eq.(1) that, if we exclude solutions in which $T_1 = 0$ and $M_1 = 0$ simultaneously, the a fixed M_3 and $T_1 = 0$ determine the solution uniquely. We will show that there are at most two orientations of the coordinate system for which $O_3 = M_3$. This will imply that (except for two special cases that will be examined separately) there are at most two distinct tile lines and therefore at most two distinct solutions. We will assume here $M_3 \neq 0$, the case $M_3 = 0$ is examined in section 4. It is convenient for the proof to assume that the velocity field satisfies initially $O_3 = O_4$. This can be assumed without loss of generality, since it is always possible to satisfy the assumption by re-orienting the coordinate system (choosing new x, y coordinates) in the following manner. Let us rotate the coordinate system along the line of sight by an angle β and denote the eight observables in the new coordiante system by $(\overline{O}_1, ..., \overline{O}_8)$. We will determine an angle β such that following the rotation $\overline{O}_3 = \overline{O}_4$. In the rotated coordinate system:

$$\begin{aligned} \overline{O}_3 &= O_3 cos^2\beta + O_4 sin^2\beta + 2O_5 sin\beta cos\beta \\ \overline{O}_4 &= O_4 cos^2\beta + O_3 sin^2\beta - 2O_5 sin\beta cos\beta \end{aligned} \qquad (5)$$

to obtain $\overline{O}_3 = \overline{O}_4$, β must satisfy

$$(O_3 - O_4)cos2\beta + 2O_5 sin2\beta = 0$$

assuming $O_5 \neq O$, β is determined by

$$tan2\beta = \frac{O_4 - O_3}{2O_5} \qquad (6)$$

There will always be a solution forβ. (Not necessarily unique. The case $O_5 = 0$ will be examined separately.) We can assume therefore that we have initially a coordinate system in which $O_3 = O_4$. We now rotate the coordinate system by a new angle α. and denote again the observables before the rotation be $(O_1, ..., O_8)$ and following the rotation by $(\bar{O}_1, ..., \bar{O}_8)$. We seek an angle α for which $\bar{O}_3 = M_3$. In general $\bar{O}_3 = O_3 cos^2\alpha + O_4\ sin^2\alpha + 2O_5 sin\ \alpha cos\ \alpha$. But since $O_3 = O_4$

$$\bar{O}_3 = O_3 + 2O_5 sin\ \alpha cos\ \alpha \tag{7}$$

assuming $O_5 \neq 0$

$$sin 2\alpha = \frac{M_3 - O_3}{O_5}$$

There are four solutions for α in the range $[0, 2\pi]$, α_1, α_2, $\alpha_1 + \pi$,$\alpha_2 + \pi$. The solution α_i,$\alpha_i + \pi$ are equivalent, they give the same solution for $(M_1, ..., M_6, T_1, T_2)$.

This establishes the claim for the general case. Two special cases that were excluded from the proof can be analyzed in a similar manner. In the first case there is a solution for which $T_1 = 0$ and $M_1 = 0$ simultaneously. In this case the velocity field will have a single direction along which $O_3 = M_3$. M_1, M_3, M_5, T_1 are then determined uniquely, but there are two solutions for M_2, M_4, T_2, in (1). In the second case $O_5 = 0$ for every orientation of the coordinate system. In this case there are two solutions, in one $T_1 = T_2 = 0$ (frontal plane) and in the other $M_1 = M_2 = 0$ (motion along the line of sight). In all of these cases the velocity field admits at most two planar interpretations.

4. Unique Solutions

In the previous section it was shown that the velocity field of a moving plane is compatible, in general, with exactly one additional moving plane.

There are two special cases under which the two-fold ambiguity disappears, and a degenerate case under which the surface orientation cannot be recovered. The discussion of these cases will be brief, since one of these cases is discussed in Ullman & Waxman (1983) and the other in Longuet-Higgins (1984).

The degenerate case arises when $M_1 = M_2 = M_3 = 0$. The motion in this case is pure rotation. The motion parameters can be recovered, but the surface orientation remains ambiguous.

One unambiguous case arises when there is no relative velocity along the observer's line of sight, i.e. $M_3 = 0$ (but M_1 and M_2 are not both zero). In this case the motion equations can be solved explicitly, and the solution is unique. The second unambiguous case arises when $M_3 \neq 0$ and the observer moves directly towards or away from the surface. If the original motion satisfies $M_1/M_3 = -T_1$, $M_2/M_3 = -T_2$ then it can be seen from equation (2) that the second solution coincides with the first. The latter condition is discussed in Longuet-Higgins, (1984). He also suggested an additional condition that can be used to resolve the ambiguity, by using an extended region of the plane rather than the local velocity field.

In the previous section we have left out the case $O_5 = 0$. Inspection of the equations reveal that this case falls into one of the categories already discussed. One case where $O_5 = 0$ is obtained when $M_1 = M_2 = M_3 = 0$, the ambiguous case discussed above. A second case is when $M_1 = M_2 = 0$ (but $M_3 \neq 0$) and $T_1 = T_2 = 0$. In this case the motion is along the surface normal, and the solution is unique. In all other cases there are two distinct solutions.

4.1 Summary

The ambiguous and non-ambiguous solution can be summarized as follows.

1. In the pure rotation case the motion can be recovered but the surface orientation remains ambiguous.
2. If there is no relative velocity along the line of sight ($M_3 = 0$ but $M_1 M_2 \neq 0$), or if the motion satisfied $M_1/M_3 = -T_1$, $M_2/M_2 = -T_2$ (motion perpendicular to the plane), then the solution is unique.
3. In all other cases the local velocity field has a two-fold ambiguity. The second solution can be derived in terms of the first by equation (2).

5. The exclusion of non-planar solutions

We have seen in the previous section that the velocity field of a moving plane has in general one additional planar interpretation. That is, if the moving surface is known to be a plane, there are in general two distinct solutions. The possibility remains, however, that when nothing is known in advance about the surface, there are additional non-planar interpretations. In this section the question of non-planar ambiguities is considered. In other words, the question is whether in addition to the two planar interpretations, non-planar solutions are also possible.

Let π be a moving plane and let μ denote the vector $(M_1, ..., M_6)$ of its motion parameters. Suppose that the twice-differentiable surface P is non-planar around the origin and that it induces the same velocity field as π, and let p denote the motion parameters of P. It is shown that if the velocity fields of P and π coincide in a neighborhood around the origin, then P is in fact planar at the origin. That is, if $\overline{T}_3, \overline{T}_4, \overline{T}_5$ are the surface curvature parameters of P at the origin then $\overline{T}_3 = \overline{T}_4 = \overline{T}_5 = 0$.

Without loss of generality we can assume that both surfaces pass through the point $(0,0, Z_o)$. We can then distinguish between two cases.

Case 1: π is tangent to P at $(0,0, Z_0)$. In this case π and P have the same values for T_1, T_2. From the original equation (1) it can be verified that the only ambiguous configuration in this case is when M_1 and M_2 in (1) both vanish. That is, the motion is directed along the line of sight. In all other cases P must coincide at the origin with π, and satisfy $\overline{T}_3 = \overline{T}_4 = \overline{T}_5 = 0$.

Case 2: π intersects P along some space-curve c. Let γ be the projection of c on the image plane. Assume first that near the origin γ is not a straight line segments. Longuet-Higgins (1984) has shown that given the image velocities of four coplanar points (no three of which are colinear in the image), the 3-D motion parameter can be recovered up to the two-fold ambiguity discussed in the previous section.

The implication is that the motion p coincides either with μ or with $\overline{\mu}$, the motion of the dual solution to μ. In either case we obtain that planar and non-planar surfaces with identical motion parameters produce an identical velocity field. From the original equation (1) it can be verified that the only ambiguous configuration in this case arises again when $M_1 = M_2 = 0$. If M_3 also vanishes, (the pure rotation case), the situation is inherently ambiguous, as discussed in the previous section. If $M_3 \neq 0$ then π is in fact tangent to P, as in the previous case.

The only remaining possibility is that near the origin γ is a straight line segment. In this case we can use the results of section 3.2. Without loss of generality γ can be assumed to lie along the x axis. Let M_i, T_i denote the motion and shape parameters of $\pi, \overline{M}_i, \overline{T}_i$ of P. From 3.2, it follows that $M_3 = \overline{M}_3, M_1 = \overline{M}_1, M_5 = \overline{M}_5$. In addition $T_1 = \overline{T}_1$ and $T_3 = \overline{T}_3$ since both are measured along c which is straight line in space common to π and P, and $\overline{T}_3 = 0$ since π is planar. $\overline{T}_4, \overline{T}_5$ can now be analyzed using eq. (1). The equality of the observable O_9 in (1) implies $M_2 T_5 = \overline{M}_2 \overline{T}_5$, and since $T_5 = 0$, $\overline{M}_2 \overline{T}_5 = 0$. Similarly O_9 and O_{12} together imply $\overline{M}_1 \overline{T}_4 = 0$, and also $M_1 \overline{T}_4 = 0$ since $\overline{M}_1 = M_1$. If $M_1 \neq 0$, $\overline{T}_4 = 0$. From O_{10} together with O_8 it now follows that $M_1 T_5 = M_1 \overline{T}_5$ and therefore $\overline{T}_5 = 0$. The case $M_1 = 0$ can be analyzed in a similar manner.

The final conclusion is that $\overline{T}_3 = \overline{T}_4 = \overline{T}_5 = 0$ except for the case $M_1 = M_2 = 0$ (motion along the line of sight).

In conclusion, we can distinguish between two cases. If the relative motion happens to be along the line of sight, then the local velocity field of a plane π is also compatible with any non-planar surface with the same relative motion parameters, and whose tangent plane coincides with π. Unlike the planar two-fold ambiguity, this ambiguity is local, and can be resolved by inspecting a larger region of the plane.

In the more general case, in which the motion component parallel to the image plane does not vanish, non-planar solutions can be ruled out, and the only remaining ambiguity is the two-fold planar ambiguity.

6. Algorithm

The proof in section 3 although not entirely constructive, leads to a possible algorithm for computing the planar solutions. The method will be outlined briefly. A different algorithm has been suggested by Longuet-Higgins (1984).

We begin by rotating the coordinate system by an angle θ to obtain $\overline{O}_3 = \overline{O}_4$ (where $\overline{O}_i$ are the observables in the rotated coordinated system). From (6) we obtain

$$tan2\theta = \frac{O_4 - O_3}{2O_5}$$

(Provided that $O_5 \neq 0$). We can therefore obtain a solution (non unique) for $sin\theta, cos\theta$. It is a straightforward computation to then compute the observables $(\overline{O}_1, ..., \overline{O}_8)$ (Waxman & Ullman, 1983).

The planar motion equations (first eight equations in (1)) can be viewed as eight linear equations in 12 unknown: $M_1, ..., M_6$, and $X_1, ..., X_6$, when $X_1 = M_1T_1$, $X_2 = M_2T_2$, $X_3 = M_1T_2$, $X_4 = M_2T_1$, $X_5 = M_3T_1$, $X_6 = M_3T_2$. The eight equations are linearly independent. Furthermore, they can be divided into four groups of two equations.

(M_3, X_1, X_2) appear only in equations (3,4).

(M_6, X_3, X_4) appear only in (5,6).

(M_1, M_5, X_5) appear only in (1,7).

(M_2, M_4, X_6) in (2,8).

As a result, each set of unknowns can be solved up to a single scalar:

$$\begin{aligned}(M_3, X_1, X_2) &= (m_3, x_1, x_2) + \alpha(\overline{m}_3, \overline{x}_1, \overline{x}_2) \\ (M_6, X_3, X_4) &= (m_6, x_3, x_4) + \beta(\overline{m}_6, \overline{x}_3, \overline{x}_4) \qquad (8) \\ (M_1, M_5, X_5) &= (m_1, m_5, x_5) + \gamma(\overline{m}_1, \overline{m}_5, \overline{x}_5) \\ (M_2, M_4, X_6) &= (m_2, m_4, x_6) + \delta(\overline{m}_2, \overline{m}_4, \overline{x}_6)\end{aligned}$$

the $m_i, \bar{m}_i, x_i, \bar{x}_i$ $(i = 1, \ldots, 6)$ are determined by solving sets of two equations in three unknowns. The scalars $\alpha, \beta, \gamma, \delta$ remain to be determined. Since $O_3 = O_4$, it follows from (1) that $M_1 T_1 = M_2 T_2$, i.e. $X_1 = X_2$, therefore $x_1 + \alpha \bar{x}_1 = x_2 - \alpha \bar{x}_2$ and hence α is determined. Two possible values for β, γ, δ are obtained from the equations.

$$\begin{aligned} X_1 X_2 &= X_3 X_4 \\ M_3 X_1 &= M_1 X_5 \\ M_3 X_2 &= M_2 X_6 \end{aligned} \tag{9}$$

respectively. Finally we choose a value for $(\alpha, \beta, \gamma, \delta)$ to satisfy the fourth independent relation $M_2 X_1 = M_1 X_4$. There will be at least one such set of values for $(\alpha, \beta, \gamma, \delta)$, and at most two. When one solution is known, the second can be found immediately using equations (2).

7. The orthographic velocity field

Previous sections have established that the parameters of a moving plane can be determined up to the two-fold ambiguity from an arbitrarily small patch near the origin. When the viewed surface patch is small, perspective effects become small, and the recovery process may become unreliable. It therefore becomes of interest to analyze the case of orthographic projection where perspective effects play no role. (In orthographic projection a space point X, Y, Z, projects to an image point $x = X, y = Y$.)

7.1 The orthographic velocity field of non-planar surfaces

Let S_1 be a non-planar surface moving in space. For the orthographic case it can be assumed that S_1 is fixed at the origin (O, O, Z_0) since the translation components are immediately recoverable. The rotation of S_1 can always be decomposed into the sum of two components: a rotation with angular velocity w (assumed to be non-zero) about an axis lying somewhere in the frontal plane (x-rotation) and a component (z-rotation) about the line of sight Z with angular velocity θ.

A second surface S_2 is said to be a *depth scaling* of S_1 if:

1. For every point (x, y, z) on S_1, (x, y, kz) is a point on S_2 for some constant k $(k \neq 0)$.
2. The rotations w_1 and w_2 are around the same axis, and $w_2 = w_1/k$.
3. $\theta_1 = \theta_2$

For non-planar surfaces the following proposition can be established. If S_1 is a possible rigid interpretation of a given orthographic velocity field, then

S_2 is another possible interpretation if and only if it is obtained from S_1 via depth scaling.

Note that if $\theta_1 - \theta_2 \neq 0$, then S_1 and S_2 have different instantaneous axes of rotation in space. The orthographic velocity field therefore does not determine uniquely the rotation axis in space.

Since our concern here is primarily with planar surfaces, the proof will be omitted, it can be found in (Ullman 1983).

7.2 The orthographic velocity field of planar surfaces

In the planar case the twofold ambiguity of planar surfaces is combined with the inherent depth-scaling ambiguity of orthographic projection. As a result the orthographic projection of a planar surface admits two interpretations, each defined up to depth scaling.

For simplicity of the analysis we can assume that in the planar velocity field all the velocity vectors are parallel to the x axis. (If the inducing object's Z-rotation is θ, then by rotating the observed velocity field by $-\theta$ all the velocity vectors will become parallel. Their direction can be taken as the x-axis.)

The velocity field $u(x, y)$ $v(x, y)$ now has the form:

$$\begin{aligned} u(x,y) &= \alpha x + \beta y \\ v(x,y) &= 0 \end{aligned} \tag{10}$$

If (w_x, w_y, w_z) is the angular velocity vector of the rotating surface (assumed to be non-zero) then:

$$\begin{aligned} w_y z - w_z y &= \alpha x + \beta y \\ w_z x - w_x z &= 0 \end{aligned} \tag{11}$$

One solution to these equations arises when $w_z = 0$. This implies $w_x = 0$ (if z is not identically zero), and $z = \frac{1}{w_y}(\alpha x + \beta y)$. This solution corresponds to a plane rotating about the vertical axis.

If $w_z \neq 0$ then $w_x \neq 0$ also and $z = \frac{w_z}{w_x} x$. This solution is also a plane, with a tilt line along the x axis.

These two possible interpretations cannot be resolved on the basis of the instantaneous velocity field. How much additional information is required to guarantee a unique solution? For non-planar objects, it can be shown (Ullman 1983) that one additional view is sufficient to remove the depth-scaling ambiguity. For planar objects, the problem is open.

The orthographic velocity field is thus inherently more ambiguous than the perspective one. Instead of two solutions there are two families of solutions, each determined up to depth scaling.

The additional ambiguity of the orthographic velocity field implies that under local analysis (i.e. using a small surface patch) the 3-D recovery process is not entirely stable. Aspects of the 3-D structure that are invariant under depth scaling are expected to be more stable than others. For planar surfaces these invariants include the orientation of the tilt line and the rotation component around the line of sight. Parameters that are not invariant under depth scaling such as surface slant are expected to be less robust.

8. Summary

1. The velocity field of a planar surface exhibits in general a two-fold ambiguity. In addition to the moving plane that has actually induced the viewed transformation there is one additional, and in general entirely different, planar solution.

2. There are special cases in which the interpretation of the local velocity field becomes unique. These cases are (i) $M_3 = 0$ (but $M_1M_2 \neq 0$), and (ii) motion directed towards or away from the surface. A degenerate case arises for pure rotation. In this case the motion parameters can be determined, but the 3-D structure remains undetermined.

3. Additional non-planar solutions are in general excluded. The exception is the case of motion directed parallel to the surface normal.

4. The two planar solutions can be computed from the eight kinematic observables, using the algorithm in section 6.

5. If one of the two planar solutions is known, the dual solution can be expressed in terms of the first using eq. 2.

6. In the orthographic case there are, instead of two solutions, two families of solutions, each determined up to depth-scaling. It is expected that for the perspective case only 3-D parameters that are invariant under depth-scaling would be robust under local analysis.

The results explored in this paper are theoretical in nature. They set some limits on the performance of any motion perceiving device. It is unknown, however, to what degree the human visual system can approach these theoretical limits. It may be of interest, therefore, to test psychophysically some of the implications of the above analysis. For example:

—Can subjects interpret the velocity field of planar surfaces? (i.e., make some reliable judgements of relative motion parameters and surface orientation). Results in (Gibson *et. al.* 1959) indicate that under some condition this is possible, although the accuracy is probably not very high.

– Can they interpret, at least to some degree, the planar velocity fields under brief presentation? If they do, how do they handle the inherent twofold ambiguity?

– Can observers interpret orthographic velocity fields? Can they recover, for example, the tilt of one or both planar solutions?

Answers to these questions may give us a better insight into the processing employed by the human visual system in the recovery of structure from motion.

Acknowledgements

I wish to thank A.Waxman, C.Longuet-Higgins, and E.Hildreth for valuable discussions and comments.

References

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.

Gibson, E.J., Gibson, J.J., Smith, O.W. & Flock.H. 1959. Motion parallax as a determinant of perceived depth. *J. Exp. Psychol.*,**8**(**1**), 40-51.

Gibson, J.J. 1950. *Perception of the Visual World.* Boston: Houghton Mifflin.

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.

Helmholtz, H. von (1925) *Treatise on Physiological Optics* (ed. J.P.C. Southall) New York: Dover.

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

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

Koenderink, J.J. & van Doorn, A.J. 1976. Local structure of movement parallax of the plane. *J.Opt. Soc. Am.*, **66**, 717-723.

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.

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

Longuet-Higgins, H.C. 1982. The role of the vertical dimension in stereoscopic vision. *Perception*, **11**, 377-386

Longuet-Higgins, H.C. 1984. The visual ambiguity of a moving plane. *Proc. Roy. Soc. Lond., B*, **223**, 165-175.

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

Miles, W.R. 1931. Movement interpretations of the silhouette of a revolving fan. *American Journal of Psychology* **43**, 392-405.

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

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-Chamgaign Coordinated Science Laboratory Report R-921.

Tsai, R.Y. & Huang, T.S. 1981. Estimating 3-D motion parameters of a rigid planar patch. Research Report R-922, Coordinated Science Laboratory, University of Illinois.

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

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

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

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

Waxman, A.M. & Ullman, S. 1983. Surface structure and 3-D motion from image flow: a kinematic analysis. Univ. of Maryland Computer Science Technical Report CS-TR-1332.

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

CS-TR Scanning Project
Document Control Form

Date : 11 / 9 / 95

Report # AIM-870

Each of the following should be identified by a checkmark:
Originating Department:

☒ Artificial Intellegence Laboratory (AI)
☐ Laboratory for Computer Science (LCS)

Document Type:

☐ Technical Report (TR) ☒ Technical Memo (TM)
☐ Other: ____________

Document Information

Number of pages: 17 (21-IMAGES)
Not to include DOD forms, printer intstructions, etc... original pages only.

Originals are:
☒ Single-sided or
☐ Double-sided

Intended to be printed as :
☐ Single-sided or
☒ Double-sided

Print type:
☐ Typewriter ☐ Offset Press ☒ Laser Print
☐ InkJet Printer ☐ Unknown ☐ Other: ____________

Check each if included with document:

☐ DOD Form ☐ Funding Agent Form ☐ Cover Page
☐ Spine ☐ Printers Notes ☐ Photo negatives
☐ Other: ____________

Page Data:

Blank Pages (by page number): ____________

Photographs/Tonal Material (by page number): ____________

Other (note description/page number):

Description : Page Number:
IMAGE MAP: (1-17) UNH'ED TITLE PAGE, 1-16
(18-21) SCANCONTROL, TRGT'S (3)

Scanning Agent Signoff:
Date Received: 11 / 9 / 95 Date Scanned: 11 / 28 / 95 Date Returned: 11 / 30 / 95

Scanning Agent Signature: Michael W. Cook

Rev 9/94 DS/LCS Document Control Form cstrform.vsd