A SIMPLE METHOD FOR RECOVERING RELATIVE DEPTH MAP IN THE CASE OF A TRANSLATING SENSOR

K. Prazdny

Computer Vision Laboratory
Computer Science Center
University of Maryland
College Park, MD 20742
USA

ABSTRACT

When an optical sensor translates along a known direction In a stationary world the relative depth map of object points on the surface can be easily obtained when some simplifying assumptions are made. Under these assumptions the angular speeds of projecting rays can be obtained from (approximations to) temporal and spatial derivatives of Image brightness. The relative depth of any two image points is then directly computable as a simple ratio of linear functions of their angular velocities because the directions of angular velocities are known in this special situation.

Keywords: Vision, image velocities, relative depth.

Although optical flows were shown to be a rich source of information, their detection poses a problem: In general, it is difficult to obtain image velocities directly from image brightness. Here we investigate a special situation where restricting assumptions are incorporated to compute the relative depth map from image motions. Prazdny [5] has shown that given two image points corresponding to two texture elements on a rigid object (or on two different objects moving rigidly togehter), their relative depth, Rd_{ij}.., defined as the ratio of their distances (S ,S.) to the center of (polar) projection, is determined by

$$(1)\quad Rd_{ij} = S_i/S_j = (\underline{A}_{ij}\cdot\bar{Q}_i)/(\underline{A}_{ij}\cdot\bar{Q}_j)$$

where $\underline{A}_{ij} = \underline{A}_i - \underline{A}_j$ is the difference of the angular velocity vectors of the projecting rays of light, and $\bar{Q}_i\bar{Q}_j$ are the unit vector specifying the visual directions along which the two linage points lie. This is a general relation which does not depend on any restriction on (relative) motion. In the special case when the observer translates along a known (but arbitrary) direction, the direction vectors of angular velocities A_i are known. In fact,

$$(2)\quad \underline{A}_i = \dot{\beta}_i(\bar{v}\times\bar{Q}_i)/\sin(\beta_i) = (v/S_i)(\bar{v}\times\bar{Q}_i)$$

so that $A_i = (v\times Q_i)/\sin(\beta_i)$ is the unit vector of $\underline{A}_i$, a n $\dot\beta_i$ s the angular speed of the ray (the magnitude of the angular velocity vector (see Figure 1). In order to apply equation (I).to recover the relative depth we have to specify $\dot\beta_i$.

In pure translation, the image elements on the projection plane (PP) move all along straight lines determined only by the position of the image points on PP and the focus of expansion (FOE) [the unique point of intersection of all these straight lines], FOE corresponds to the point where the vector along which the observer translates pierces the PP. In the case of pure translation the directions of the image velocities are independent of the surface layout; the information about the spatial disposition of the texture elements is contained only in the magnitudes of the image motions. Because each image element moves along a known straight line on PP, the problem reduces to estimating the motion of a one dimensional brightness distribution along the line (it is assumed here that the distribution does not change significantly from frame to frame). This problem is easily solvable. Using the Taylor series expansion one can develop an expression for the image displacement A£ relating the temporal and spatial derivative of the image brightness function along the line it at a given image point (see e.g. [6]).

$$(3)\quad \Delta I = (d^2I/d\ell^2)(\Delta\ell^2/2) - (dI/d\ell)\Delta\ell$$

where I is the image brightness function, and A£ is its displacement at an image point along the line *i.* With a sufficiently small interframe interval, $\Delta I = dI/dt$, and $\Delta\ell = d\ell/dt$. If one disregards the components of order 2 and higher, this approximation reduces to a linear approximation used e.g. in [1],(2),[3],[4]. Observe that the linear approximation is appropriate only at points where the first spatial derivative along I is (approximately) constant and the displacement remains within the extent of this linearity. The speed dl/dt is related to the angular speed $d\beta/dt$ by

$$(4)\quad d\beta/dt = -(dI/dt)(d\ell/dI)(\sin(\lambda)/Q) = I/(dI/d\ell)\ \sin(\lambda)/Q$$

where Q is the distance of the image element on PP from the center of projection, and λ is the angle between the line I on PP and the projection ray (Figure 1). Q and λ are constants for a given FOE and reflect the fact that the projection surface is a plane which is not isotropic with respect to the image velocities. The above equations enable one to compute the relative depth directly. The derivatives dl/dt, and dl/dil have to be estimated from a succession of digital images. The derivative $dI/dt = I$ can be approximated directly from successive (temporally proximal) digital images. The derivative dl/dl of I along l is easily obtainable from neighboring gray level values by interpolation. One has to assure the existence of both derivatives at locations where B is to be computed, i.e. the image brightness function has to be differentiable at all such points. To do this (and to suppress the noise), the individual image is low-pass filtered before estimating the derivatives. This operation virtually guarantees the differentiability of the image brightness function nearly everywhere. The requirement on the smoothness of the image brightness function is the strength and the weakness of the method (see also [6]). On the one hand, it enables one to compute image motion from equation (3). On the other hand, it is limited to areas of large brightness changes. Observe that the image motions resolvable by the method must be small with respect to the (local) variations in the image intensity. This requirement will be satisfied near places where the gradients are large in magnitude and extent (typically, the edges).

One appealing aspect of the relative depth map as a representational structure is its potential redundancy. For a given image point, we can compute a set of relative depth values pairwise with other image points. While these values cannot be directly compared, the absolute depth of a given point has to be unique. Setting the "absolute" depth of an arbitrary point to an (arbitrary) value and propagating we obtain many independent estimates of "absolute" depth of a given point. These values can be treated as noisy estimates of the true values. One can thus devise an iterative network which not only suppresses the noise but also incorporates the constraint on smoothness of the "absolute" depth values to achieve a (global) consistency.

DISCUSSION AND CONCLUSION

The present scheme is based on the assumption that the observer purely translates and that FOE is known. This restriction reduces a difficult (and still unsolved) 2D problem to an easy 1D problem. The case of pure translation is important in many industrial applications where the camera motion can be restricted and the orientation of the optical axis of the camera with respect to some fixed vector can be determined with high precision. The method presented here should be reasonably successful in determining the relative depth between the places of large brightness variation. It is not intended for accurate computations of local surface orientation. Rather, it is expected to be useful in determining the approximate relative positions of objects (e.g. for hand-eye coordination), and in determining the occlusion relations. The method is currently being tested (using a TV camera) on real world data. It is expected that it will eventually perform in real time when implemented on dedicated hardware.

REFERENCES

[1] Cafforio, C. and Rocca, F., 1976, Methods for measuring small displacements of television images, IEEE Trans. Inf. Theory, 5, 573-579.

[2] Fennema, C. L. and Thompson, W. B. 1979, Velocity determination in scenes containing several moving objects, Comp. Graph. Image Proc. 9, 301-315.

[3] Horn, B. K. P., and Schunck, B. G. 1980, Determining optical flow, AI MIT Memo No. 572 (April 1980).

[4] Limb, J. O., and Murphy, J. A., 1975, Measuring the speed of moving objects from television signals, IEEE Trans. Commun. 23, 474-478.

[5] Prazdny, K. 1981, Relative depth and local surface orientation from image motions, Proceedings DARPA Image Understanding Workshop Science Applications Inc., McLean, VA.

[6] Snyder, W. E., Rajala, S. A. and Hirzinger, G. 1980, Image modelling, the continuity assumption and tracking, IEEE Conf. on Patt. Rec. and Image Proc., Miami Beach, FL, December 1980.

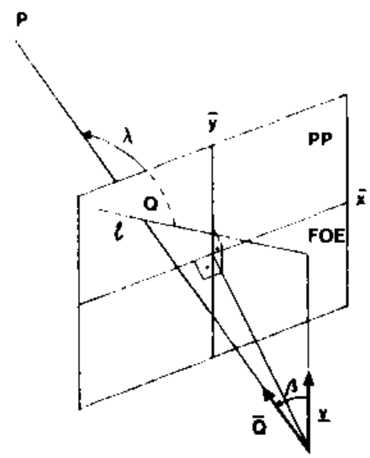

Figure 1