Correspondence

Obstacle Avoidance Using Flow Field Divergence

RANDAL C. NELSON AND JOHN (YIANNIS) ALOIMONOS

Abstract—**The use of certain measures of flow field divergence is investigated as a qualitative cue for obstacle avoidance during visual navigation. It is shown that a quantity termed the *directional divergence* of the 2D motion field can be used as a reliable indicator of the presence of obstacles in the visual field of an observer undergoing generalized rotational and translational motion. Moreover, the necessary measurements can be robustly obtained from real image sequences. Experimental results are presented showing that the system responds as expected to divergence in real world image sequences, and the use of the system to navigate between obstacles is demonstrated.**

Index Terms—**Image flow, motion analysis, motion detection, visual navigation.**

I. Introduction

Navigation is a basic operation in automation which must be performed by any system that manipulates or responds to physical objects within its environment. Vision has received particular attention in this connection, primarily because of its potential to provide a large amount of information about many different characteristics of the environment. Much vision research has sought to provide general theoretical frameworks from which a broad range of problems could be addressed. However, in order to achieve the desired generality, such approaches frequently make assumptions which are inappropriate in real-world environments. An alternative approach is to investigate specific, well-defined problems having general utility, and attempt to provide a solution that can be both adequately described theoretically, and demonstrated to work in real-world environments. This recalls a program advocated by Brooks [4] who proposes to understand intelligent behavior by duplicating progressively more sophisticated abilities displayed by living organisms.

One of the most elementary forms of navigation is obstacle avoidance by a moving, compact sensor. It is a prerequisite, however, for many more complex abilities since any system performing a more complicated task must avoid obstacles in the process, and is thus one specific problem for which a general solution is highly desirable. We argue that the flow field divergence is a particularly useful cue for this problem, and one that can be robustly derived from image sequences in real-world environments.

II. Flow Field Divergence as a Qualitative Motion Cue

A camera moving within a three dimensional environment produces a time-varying image which can be characterized at any time t by a two-dimensional vector-valued function f known as the *motion field*. The motion field describes the two dimensional projection of the three-dimensional motion of scene points relative to the camera. In general, the motion field depends on the motion of the camera, the structure of the environment, and the motion (if any) of objects in the environment. If all these components are known, then it is straightforward to calculate the motion field.

For machine vision, it would be useful to invert the process to obtain information about camera motion and structure of the environment from the motion field. In general, this problem is ill-conditioned, and additional assumptions about smoothness and rigidity are necessary to obtain a unique solution. There is a large body of work on the theory of extracting shape and/or motion information from both point correspondences [6], [14], [15] and dense motion fields [3], [17], [12]. Most of these studies, however, make the assumption that accurate information is available. Unfortunately, the solutions to the equations are frequently inordinately sensitive to small errors in the motion field [14]. Neither point correspondence methods [2], [7], [8] nor differential methods aimed at extracting dense flow fields [1], [5], [9], [18] have yielded information of sufficient accuracy to allow the theoretical results to be reliably applied in practice.

We propose to make use of qualitative motion features which can be robustly determined to the necessary accuracy using simple differential methods. There are two basic sources of error in such methods. The first is that the local apparent motion of the image, known as the *optical flow*, does not necessarily correspond to the 2D motion field. For instance, a spinning, featureless sphere under constant illumination has zero optical flow, but a nonzero motion field. Verri and Poggio [16] have shown that the motion field and the optical flow correspond exactly only under special conditions of lighting and movement. They also show, however, that for sufficiently high gradient magnitude, the agreement becomes arbitrarily close. Thus for strongly textured images the motion field and the optical flow are approximately equal. A second problem with differential techniques is that only the component of the optical flow parallel to the local image gradient can be recovered locally. This is known as the *aperture problem* and it corresponds to the intuitive observation that for a moving edge, only the component of motion perpendicular to the edge can be determined. It turns out that, for the problem of obstacle avoidance, the flow information can be used in projected form thus sidestepping the problem.

The qualitative measurements we propose for obstacle avoidance are based on the divergence of the motion field. Possible uses for divergence are mentioned briefly by Thompson [13], but not developed. The obvious motivation stems from the fact that an obstacle in relative motion towards the camera produces an expanding image, i.e., one whose image flow has positive divergence. Thus, regions of positive divergence represent potential obstacles with the distance to the obstacle inversely proportional to the magnitude of the divergence. It can also be shown that divergence is invariant under rotational motion of the sensor which is a valuable property because it allows the cue to be utilized even when the sensor is not completely stabilized. Approaching obstacles do not represent the only situation producing positive divergence in the motion field. A tilted surface translating parallel to the image plane produces divergent flow, and certain motion discontinuities are also interpretable as divergence. On the other hand, divergence is always associated with a nearby object. The remainder of this section is devoted to an analysis of this relationship.

We consider the image formed by spherical projection of the environment onto a sphere of radius ρ termed the *image sphere*. The use of spherical projection makes all points in the image geometrically equivalent, which simplifies the analyses. Ordinary cameras do not utilize spherical projection, but if the field of view is not too wide, the approximation is reasonably close. Since the distortion is purely geometric in origin, it could be corrected if

Manuscript received September 8, 1987; revised January 30, 1989. Recommended for acceptance by W. Thompson.

R. C. Nelson is with the Department of Computer Science, University of Rochester, Rochester, NY 14627.

J. Y. Aloimonos is with the University of Maryland, College Park, MD 20742.

IEEE Log Number 8929342.

necessary. In experiments we performed using a camera with a field of view approximately 20 × 30 degrees, no correction was necessary to obtain usable results.

For collision avoidance, the components of relative motion parallel to and perpendicular to the sensor are of primary interest. For a given point p on the image sphere, these can be conveniently expressed in terms of local coordinate systems defined by p and the center of projection c. Points in the environment are specified by coordinates (X, Y, Z) where $(0, 0, 0)$ coincides with the center of projection, and the positive Z axis passes through p. Image positions in the neighborhood of p are specified by coordinates (x, y) where $(0, 0)$ coincides with p and the x and y axes with the local projection of the X and Y axes, respectively. This is permissible because the image sphere is locally Euclidean. The Euclidean neighborhood of p will be referred to as the *local projective plane*. Since all points in the image are geometrically equivalent under spherical projection, and the divergence is a local operation, all of our analysis can be done in terms of these local coordinate systems.

The relevant characteristics of the object projecting to p are its distance from the center of projection and the orientation of the surface. Since the divergence involves only first order derivatives, higher order parameters are not significant. The distance to the object is just its Z coordinate. The orientation of the surface can be described by the angles α and β in gradient space where the *tilt*, α, is the angle of the projection of the surface normal on the image sphere, and the *slant*, β is the angle between the surface normal and the Z axis. This geometry is shown in Fig. 1.

We now define a parameterized, one-dimensional divergence measurement and derive the relationships between this measurement and motion of the sensor relative to objects in the environment. This parameterized version of the divergence both contains more information than the usual scalar, and is more directly calculable from real image sequences.

Let f be the motion field under spherical projection. We define the *directional divergence* $D_\phi f$ to be the one-dimensional divergence of f in the direction ϕ. Symbolically, we write

$$D_\phi f = \frac{\partial f_\phi}{\partial r_\phi}$$

where f_ϕ is the component of f in the ϕ direction and r_ϕ is Euclidean distance in direction ϕ. In terms of the local x–y coordinate system, it is easy to show that

$$D_\phi f = \cos^2 \phi \frac{\partial f_x}{\partial x} + \sin^2 \phi \frac{\partial f_y}{\partial y} + \sin \phi \cos \phi \left[\frac{\partial f_x}{\partial y} + \frac{\partial f_y}{\partial x} \right]$$

where ϕ is measured from the x axis.

Clearly, D_ϕ is a linear operator. Thus we can separate the contributions of different motions and consider them independently. For motion in a rigid, three-dimensional environment, the motion field at any point p on the image sphere can be represented as the superposition of three separate motion fields: one due to sensor rotation with magnitude $\vec{\omega}$, one due to translation perpendicular to the image sphere at point p with magnitude v_{perp}, and one due to translation parallel to the image sphere at p with magnitude v_{par}. We will refer to these motion fields as f_{rot}, f_{perp}, and f_{par}, respectively. Hence,

$$f = f_{\text{rot}} + f_{\text{perp}} + f_{\text{par}},$$

and since D_ϕ is linear

$$D_\phi f = D_\phi f_{\text{rot}} + D_\phi f_{\text{perp}} + D_\phi f_{\text{par}}.$$

The following three equations describe the relationship between the various types of motion and the directional divergence. Derivations can be found in [11]. For rotational motion, we have

$$D_\phi f_{\text{rot}} \equiv 0.$$

In other words, directional divergence is invariant under rotational motion. Thus its use as a cue is not limited to situations where the sensor is undergoing purely translational motion. For translation perpendicular to the sensor we have

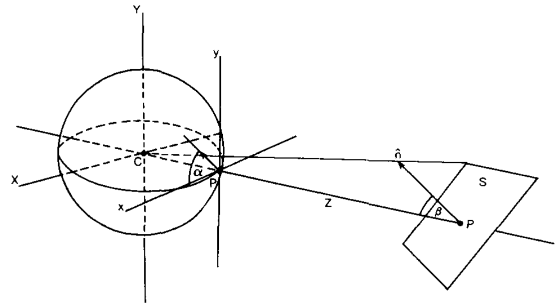

Fig. 1. Spherical projection geometry and local coordinate systems.

$$D_\phi f_{\text{perp}} = \frac{v_{\text{perp}}}{Z},$$

independent of ϕ and the orientation of the surface. This divergence corresponds to the apparent expansion of an approaching object. Translation parallel to an obstacle produces no divergence if the surface is parallel to the direction of translation, in which case locally constant flow is produced. However, if the surface is tilted, the changing depth produces divergence in the motion field. The effect is given by

$$D_\phi f_{\text{par}} = \frac{\tan \beta v_{\text{par}}}{Z} [\cos^2 \phi \cos \alpha + \cos \phi \sin \phi \sin \alpha].$$

Since the component of translation parallel to an obstacle can produce divergence in the motion field, the presence of divergence in the flow is not simply equivalent to the presence of an obstacle on a collision course with the sensor. However, by the above formula, the directional divergence due to f_{par} is zero when $\cos \phi = 0$, that is, when ϕ is perpendicular to the direction of parallel translation. Thus if the distance between the sensor and an object projecting to point p on the image sphere is decreasing, there exists a direction ϕ for which the directional divergence $D_\phi f(p)$ is positive. Thus, any object on a collision course with the sensor can be detected by testing whether there exists some ϕ for which D_ϕ is positive.

Not all positive divergence indicates an imminent collision. As we have seen, translation parallel to a inclined surface can produce divergence flow. In addition, even if an object has a perpendicular component to its motion, it may also have a horizontal component that will prevent a collision. The presence of the factor $1/Z$ in all the terms which contribute to the divergence does suggest, however, that divergence indicates the proximity of an object. The presence of the term $\tan \beta$ in the expression for the divergence due to parallel motion would seem to complicate matters since it becomes arbitrarily large as β approaches 90°, corresponding to a depth discontinuity. It can be shown however, that the total change in the flow is bounded. In particular, the following can be shown. If $|D_\phi| \geq D > 0$ within a disk of radius r about p where r is small with respect to the radius ρ of the image sphere, then

$$Z_{\min} \leq \max \left\{ \left| \frac{2v_{\text{perp}}}{D} \right|, \left| \frac{\rho v_{\text{par}}}{rD} \right| \right\}$$

where $Z_{\min}$ is the distance to the nearest point projecting into the disk.

Thus the divergence due to a relatively distant object can be large, but only over a very short distance in the image. The presence of divergence over any significant area thus indicates an object which is nearby in terms of the translational velocity of the sensor.

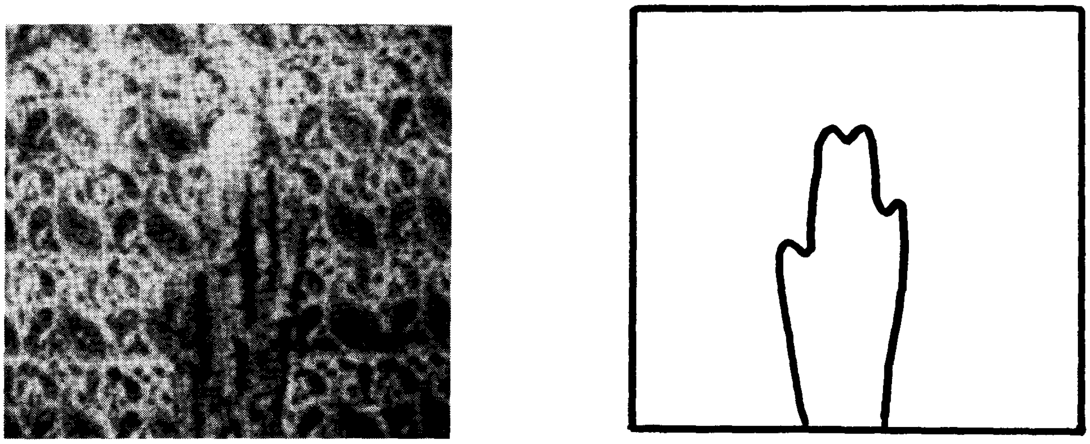

Fig. 2. Obstacle (piece of bark) against textured background.

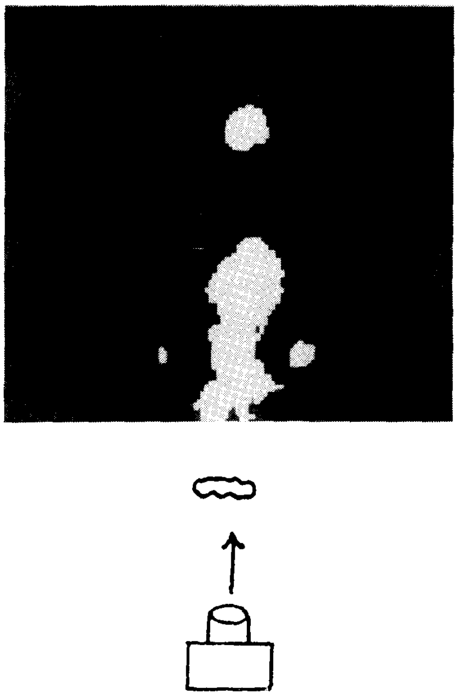

Fig. 3. Perpendicular translation.

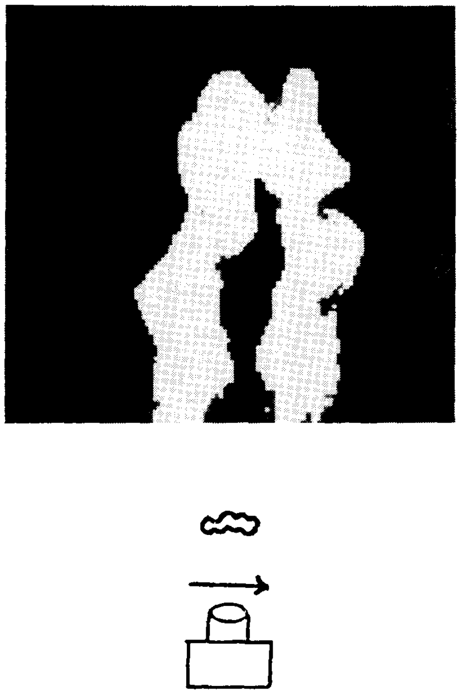

Fig. 4. Parallel translation.

The theoretical results can be summarized as follows. For a sensor with arbitrary rotational and translational motion, any object on a collision course with the sensor will produce a positive directional divergence in the motion field for some angle ϕ; however, not all positive divergence indicates an imminent collision. Conversely, any nonzero directional divergence measurement indicates the presence of an object which is nearby in terms of the relative translational velocities of the sensor and the object; however, not all nearby objects produce divergence in the motion field. These results are sufficient for low-level obstacle avoidance, although used alone, they could result in maneuvers which are not strictly necessary to avoid collision.

The results can be strengthened if only translational motion is permitted. In this case we have the well known result that the object projecting to p is on a collision course if and only if D_ϕ is positive at p, and the flow is zero (see, e.g., [13]). p then corresponds to a focus of expansion. If it were possible to stabilize the sensor, either by inertial means or by a visual algorithm (see [10]), this could be quite useful.

The results obtained above suggest that detection of directional divergence in the motion field could form the basis for a simple visual obstacle avoidance system. It turns out that the required divergence measures can be obtained robustly from real image sequences using simple, differential techniques. Use of a differential method means dealing both with the nonidentity of the motion field with the optical flow, and with the aperture problem. The first problem was eliminated by assuming enough visual texture in the image to avoid any difficulty. Ordinary natural objects such as tree

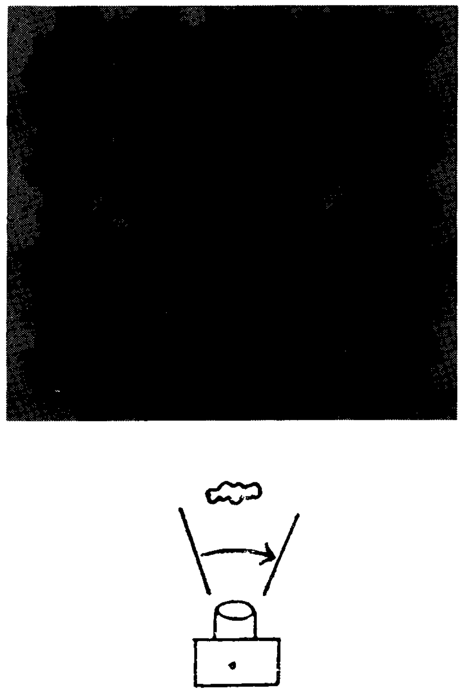

Fig. 5. Rotation (pan).

Fig. 6. Rotation (roll).

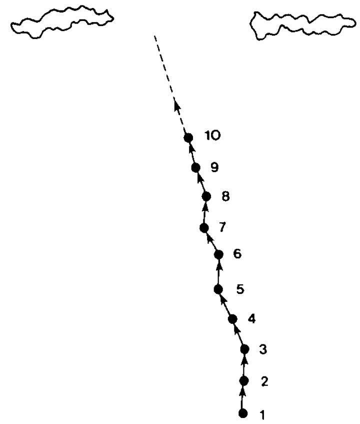

Fig. 7. Path of robot navigating between a pair of obstacles.

bark, broken brick, and human faces were found to have sufficient visual texture to make the method workable. The aperture problem can be circumvented because only projected flow information is needed to compute the directional divergences. For each of several principal directions, the parallel flow component is approximated at just those points in the image where the gradient happens to be in the same direction. By accumulating this sparse, approximate data over time using many images, enough information is obtained to compute the directional divergences for each principal direction. The internal thresholds used to determine whether a value represents information or noise were determined *a priori* from geometric error analysis. Hence, no extensive testing was necessary in order to determine parameter values. Details can be found in [11].

III. Implementation and Testing

The utility of divergence information was investigated empirically using a robot arm to move a miniature CCD camera through a three dimensional environment containing various obstacles. We first tested the response of the system to individual translational and rotational movements to see whether the output of the divergence detector corresponded to theoretical predictions. Fig. 2 shows the test scene as viewed by the camera. The scene consists of an obstacle, in this case, a piece of bark, in front of a distant background. The first test consisted of moving the camera directly forward towards the obstacle. Fig. 3 shows the hazard map generated by motion directly towards the obstacle. White areas correspond to regions labeled "danger," gray to regions labeled "caution," and black to "safe" regions. Theoretically, the obstacle should display a greater divergence than the background and should be detectable on that basis. The bark stands out clearly as an obstacle. Next, the camera was moved parallel to the obstacle. In this case, the flow discontinuities at the boundaries of the obstacle should stand out. Fig. 4 shows that this is the case. We next tested the invariance under rotation. According to the analysis, the image sequence resulting from pan, tilt, or roll of the camera should show no divergence. Fig. 5 shows the hazard map resulting from camera pan. As expected, little or no divergence is detected. Fig. 6 shows the hazard map resulting from camera roll. The flow field for this motion is a circular whirlpool about the center of the image. There are a few small regions of spurious divergence reported for this sequence, but basically, the map is clear as expected.

The above tests indicated that the divergence detector worked fairly robustly in detecting the qualitative divergence features predicted by the theoretical analysis. The information should thus be usable for navigational purposes. We demonstrated this by implementing a simple navigation algorithm. Basically, the camera moves straight ahead for a short distance, acquiring a sequence of images which is analyzed to produce a hazard map. The map is divided into sections by a 3×3 grid, and each section is assigned a hazard index by adding the number of "danger" points to $1/10$ the number of "caution" points. The direction of movement is

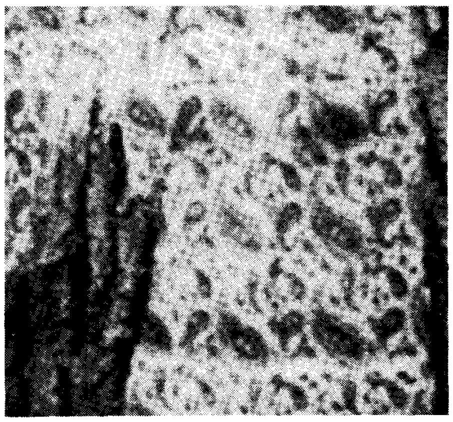

Fig. 8. Scene at step 4.

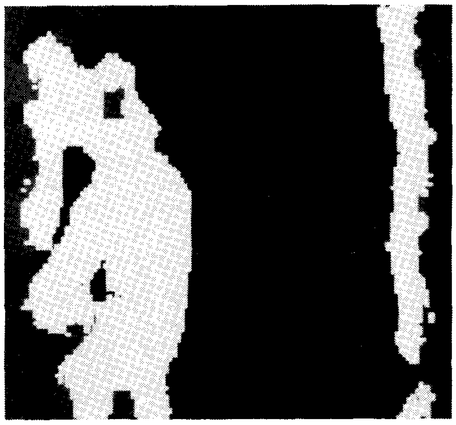

Fig. 9. Hazard map at step 4.

then updated to point towards the center of the section with the lowest hazard index, and the process repeats.

Fig. 7 shows the results of a test run where the system navigated through the gap between two pieces of bark. The obstacles were large enough to prevent the robot from seeing around them, and it was prevented from navigating up and over the obstacles by allowing only left-right directional changes. This forced the system to find the gap. The figure shows the path followed by the camera as it approached the obstacles. Note how the robot switched course back and forth to in order to keep heading for the gap between the obstacles. Figs. 8 and 9 show the scene viewed by the camera and the generated hazard map at step 4.

IV. Conclusions

We have argued that the flow field divergence represents a qualitative measurement which is useful for obstacle avoidance during visual navigation. We have further demonstrated that the flow field divergence can be computed robustly from real image sequences using relatively simple operations. Two basic principles which facilitated this task were directional decomposition of the problem into multiple one-dimensional domains, and temporal accumulation of information which permitted the use of many-image sequences. Finally, we presented a practical demonstration in which image flow divergence was used to maneuver a vehicle between obstacles in a real world environment. The technique has applications in the visual guidance of industrial robots, in the navigation of autonomous aircraft or submarine vehicles, and as a fundamental component of more general vision systems.

References

[1] P. Anandan and R. Weiss, "Introducing a smoothness constraint in a matching approach for the computation of optical flow fields," in *Proc. Third Workshop Comput. Vision: Representation Contr.*, 1985, pp. 186–194.

[2] S. Barnard and W. Thompson, "Disparity analysis of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-2, pp. 333–340, 1980.

[3] R. C. Bolles, "Epipolar plane analysis: An approach to determining structure from motion," *Proc. IJCAI*, pp. 7–15, 1987.

[4] R. A. Brooks, "Achieving artificial intelligence through building robots," A. I. Memo 899, *Massachusetts Inst. Technol. Artificial Intell. Lab.*, May 1986.

[5] B. K. P. Horn and B. G. Schunk, "Determining optical flow," *Artificial Intell.*, vol. 17, pp. 185–204, 1981.

[6] H. C. Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections," *Nature*, vol. 293, 1981.

[7] D. Marr and T. Poggio, "A computational theory of human stereo vision," *Proc. Roy. Soc. London*, vol. B 204, pp. 301–328, 1979.

[8] H. P. Morovec, "Visual mapping by a robot rover," *Proc. IJCAI*, pp. 598–600, 1979.

[9] H. H. Nagel and W. Enkelmann, "An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-85, pp. 565–593, Sept. 1986.

[10] R. C. Nelson and J. Aloimonos, "Finding motion parameters from spherical flow fields (or the advantages of having eyes in the back of your head)," *Biological Cybern.*, vol. 58, pp. 261–273, 1988.

[11] ——, "Using flow field divergence for obstacle avoidance in visual navigation," in *Proc. DARPA Image Understanding Workshop*, Apr. 1988, pp. 548–567; also Univ. of Maryland Comput. Sci. TR #1840.

[12] K. Prazdny, "Determining the instantaneous direction of motion from optical flow generated by a curvilinear moving observer," *Comput. Vision Graph. Image Processing*, vol. 22, pp. 238–248, 1981.

[13] W. B. Thompson and J. K. Kearney, "Inexact vision," in *Proc. Workshop Motion, Representation, Anal.*, May 1986, pp. 15–22.

[14] R. Y. Tsai and T. S. Huang, "Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces," *IEEE Trans. Pattern Anal., Machine Intell.*, vol. PAMI-6, pp. 13–27, 1984.

[15] S. Ullman, "The interpretation of structure from motion," *Proc. Roy. Soc. London*, vol. B 203, pp. 405–426, 1979.

[16] A. Verri and T. Poggio, "Against quantitative optical flow," *Int. Conf. Comput. Vision*, June 1987, pp. 171–180.

[17] A. Waxman, "Image flow theory: A framework for 3-D inference from time varying imagery," in *Advances in Computer Vision*, C. Brown, Ed. Lawrence Erlbaum Inc. 1987.

[18] Y. Yasumoto, "Experiments in estimation of 3-D motion parameters from a sequence of image frames," in *Proc. 1985 IEEE Comput. Soc. Conf. Comput. Vision Pattern Recog.*, pp. 560–671.

Localization and Noise in Edge Detection

E. De Micheli, B. Caprile, P. Ottonello, and V. Torre

***Abstract*—In this correspondence, we analyze two aspects of edge detection: accuracy of localization and sensitivity to noise. The detection of corners and trihedral vertices is analyzed for gradient schemes and zero-crossing schemes. Neither scheme correctly detects corners or trihedral vertices, but the gradient schemes are less sensitive to noise.**

A simple but important conclusion of this correspondence is that the noise present in digital images of typical indoor scenes is small and the signal-to-noise ratio is high. In fact, the noise present in digital images

Manuscript received March 7, 1988; revised November 23, 1988. Recommended for acceptance by W. E. L. Grimson.

The authors are with the Department of Physics, Via Dodecaneso 33, 16146 Genova, Italy.

IEEE Log Number 8929344.

0162-8828/89/1000-1106$01.00