Correspondence

A Diffusion Mechanism for Obstacle Detection from Size-Change Information

Dario L. Ringach and Yoram Baram

Abstract—A mechanism for the visual detection of obstacles is presented. A new immediacy measure, representing the imminence of collision between an object and a moving observer, is defined. A diffusion process on the image domain, whose initial condition is determined by the motion field normal to the object's boundary, is shown to converge asymptotically to the immediacy measure. A network of locally connected cells, derived from a finite-difference approximation of the diffusion equation, estimates the immediacy measure from normal velocity and boundary information provided by a motion measurement and segmentation stage. The algorithm's performance on real image sequences is demonstrated.

Index Terms—Obstacle detection, collision avoidance, motion field, immediacy measure, diffusion process, size-change information.

I. INTRODUCTION

Obstacle detection is a basic task in visual navigation. Early approaches to visual motion problems have been based on the attempt to recover the exact shapes of the objects in the scene and their relative velocity with respect to the observer from motion measurements [1]-[7]. The resulting mechanisms require highly constrained models and are sensitive to measurement errors [8]-[11]. These problems can be largely circumvented by realizing that performing the task of obstacle detection does not really require accurate estimates of their shapes and velocities. The approximate location of the objects involved and some qualitative measure of the imminence of collision would normally suffice. Inexact solutions to motion processing problems have been proposed in the past [12], [13].

When an observer approaches a textured object, the object's image, defined by its projection onto a small sphere (retina, camera sensor) centered at the observer's location, grows larger and the granularity of its surface texture grows coarser. These two cues are called size-change and texture-change, respectively [14]. There are situations in which at least one of these cues is absent. In moving towards a large object, for instance, size-change information may vanish as the object's boundaries drop out of the image domain. On the other hand, many objects do not possess a textured exterior surface. In a recent paper, Nelson and Aloimonos [13] based an obstacle avoidance system on motion field divergence-like measurements, derived from texture-change information. Their underlying assumption is that the objects, as well as the background, are highly textured, so that their differential measure can be estimated almost everywhere in the image domain.

In this note, we propose a method for obstacle detection from size-change information. Our approach may be viewed, then, as complementing that of Nelson and Aloimonos. A new immediacy measure, representing the imminence of collision between an object and a moving observer, is defined. A diffusion process on the image domain, whose initial condition is determined by the component of the motion field normal to the object's boundary, is shown to converge asymptotically to the immediacy measure. Bounds on the convergence time are described. Normal velocity and object boundary information is provided by simple motion measurement and edge detection techniques. A network of locally connected formal neurons, estimating the immediacy measure, is derived from a finite-difference approximation of the continuous diffusion equation. The proposed mechanism, which is particularly suitable for parallel hardware implementation, has been simulated on a sequential computer and shown to be effective in the detection of real objects moving towards a camera. A more detailed account of this work, including bounds on the convergence time of the diffusion process and fault tolerance analysis, can be found in [50].

II. THE OBJECT IMMEDIACY MEASURE

The first step in approaching the problem at hand is defining a measure by which objects in the image domain can be rated as obstacles. The geometry of the problem is illustrated in Fig. 1. Consider the image formed by the projection of the environment onto the unit sphere $S$—called the image sphere—with center at the point $O$. Let $P$ be the projection of a point $P$ in the three dimensional space onto $S$, which is defined as the intersection of $S$ with the semi-infinite ray with endpoint at $O$ passing through $P$. Let $(X, Y, Z)$ be a Cartesian coordinate system with origin at $O$ such that $P$ lies on the positive $Z$ axis. We define the local projective image plane of $P$ to be the tangent plane to $S$ at $P$, and attach to it a local coordinate system $(x, y, z)$ with center at $P$ obtained by the projection of the $(X, Y)$ coordinate system onto the local projective plane of $P$. Suppose that $P$ is on some smooth (having first order derivatives) rigid surface described by $Z = f(X, Y)$. We assume that the relative motion of a camera, located at $O$, with respect to the surface is given by a translational velocity $V = (V_x, V_y, V_z)$ and a rotational velocity $\Omega = (\Omega_x, \Omega_y, \Omega_z)$.

0162-8828/94$04.00 © 1994 IEEE
The motion field is defined as the vector field on the image sphere obtained by projecting the 3-D velocity field of the visible surfaces of the objects in the scene onto S. We denote by $v(p) = (u,v)$ the motion field at $p$, expressed in the local coordinate system $(x,y)$ of $p$. The local motion field in the neighborhood of $p$ can now be analyzed at the local image plane of $p$. Let $Z = Z_0 + VZ$ be the first order approximation of the surface in the neighborhood of the origin, where $Z_0$ is the distance from $O$ to $P$, and $\nabla Z = (Z_X, Z_Y)$ is the surface gradient. We denote by $(V_x, V_y, V_z) = (V_x, V_y, V_z)/Z_0$ the normalized translational velocity. The following relationship between the surface parameters, the motion of the camera, and the motion field at $p$ holds [2]:

$$
u = -V_y - \Omega X, \quad v = -V_x + \Omega Y,$$

$$u = V_x + V_z Z_Y, \quad v = V_x + V_z Z_X,$$

$$w = -V_z Z_Y + V_x Z_Y, \quad v = -V_z Z_X + V_x Z_X, \quad (1)$$

where $(u,v)$ is the image velocity at $p$, and $u_x, u_y, v_x, v_y$ are its partial derivatives there.

As a consequence of the relative motion, the projection of an infinitesimal surface patch around $P$ onto the image sphere defines a neighborhood of $p$ which deforms with time. We define the immediacy of a point $p$ on the image sphere, which we denote by $\psi(p)$, as the relative rate of expansion $(\psi(p) > 0)$ or contraction $(\psi(p) < 0)$ of an infinitesimal image neighborhood of $p$. This quantity is given by the divergence of the motion field at $p$ [16], [17]:

$$\psi(p) = \nabla \cdot (u,v), \quad (2)$$

yielding, by (1),

$$\psi(p) = 2V + (V_x, V_y) \cdot \nabla Z. \quad (3)$$

Suppose that $(V_x, V_y) = 0$, i.e., the movement is in the $Z$ direction. In this case, the normalized velocity $V_z = V_z/Z_0$ is the reciprocal of the time needed for the surface patch at distance $Z_0$ to collide with the center of the camera, $O$. This measure has been termed time-to-collision inverse or "immediacy" [18]. We see from (3) that $\psi(p)$ is proportional to the time-to-collision inverse in this case. If $(V_x, V_y) \neq 0$ and $V_z = 0$, there can be relative depth changes between the surface patch and the center of the image sphere caused by the fact that the camera moves in parallel to a slanted surface. These changes are expressed by the second term in (3), which, as can be easily seen, represents the inverse of the time required for the center of the camera to collide with the plane tangent to the surface at $P$. While certain works suggest the use of upper and lower bounds on $V_z$ alone as the basis for obstacle detection [19], [20], we argue that relative depth changes caused by parallel motion should be taken into consideration in the definition of the immediacy of points, as they also embody a potential for collision.

Suppose that $R$ is the projection of an object in the three dimensional space onto the image sphere $S$. We assume the region $R$ to be a connected and closed subset of $S$, having a piecewise differentiable boundary $\partial R$. For each interior point $p \in R \setminus \partial R$ an immediacy measure $\psi(p)$ was already defined (on $\partial R$ the motion field is not continuous in general). Let

$$\chi(R) = \frac{1}{A(R)} \int_{\partial R} \psi(p) \, ds = \frac{1}{A(R)} \int_{\partial R} \nabla \cdot \psi(p) \, ds, \quad (4)$$

where $A(R)$ is the area of the region $R$, be the immediacy of the object whose projection onto $S$ is $R$. The immediacy of an object is defined, then, by the average immediacy of the points in its projection domain $R$ or, equivalently, by the average divergence of the motion field over this domain. By means of the divergence theorem [21], we obtain

$$\chi(R) = \frac{1}{A(R)} \int_R \nabla \cdot \psi(p) \cdot n \, dt, \quad (5)$$

where $n$ is the exterior unit vector normal to $\partial R$, and $dl$ is an infinitesimal arc length element of $\partial R$. The latter result means that the object immediacy can be calculated, in principle, from the component of the velocities normal to the boundary. The main advantage of defining $\chi(R)$ by (4) is that its computation does not require estimating the motion field and its divergence, which are ill-posed problems [22], [23]. It can be readily verified that $\chi(R)$ is the relative rate of variation of $A(R)$ [24]:

$$\chi(R) = -\frac{1}{A(R)} \frac{dA(R)}{dt}, \quad (6)$$

which relates the immediacy measure to size-change information.

An alternative definition of the object immediacy measure is to take the maximum of $\psi(p)$ over the region $R$. This corresponds to the intuitive idea of the object immediacy being equal to that of the most "dangerous" point in its projection. This definition would require the computation of the full motion field and its divergence. As a result, the estimated measure can be expected to be highly sensitive to noise.

We suggest $\chi(R)$ as a measure of the imminence of collision between a moving observer and an object. This definition is consistent with certain works on motion control in humans as well as in lower animals. It has been suggested, for example, that pre-landing deceleration of flying insects is determined by the relative retinal expansion velocity [25], which is the same as $\chi(R)$. In [26], a theory is presented on how a driver might visually control braking, based on time-to-collision information. In [27], it was shown that time-to-collision can be perceived by a human observer in the absence of distance and velocity information. Recently, it was suggested that relative retinal expansion velocity may also explain the head-bobbing of pigeons [28], and the chasing behavior of small crustaceans [29].

III. CALCULATION OF OBJECT IMMEDIACY BY Diffusion

Consider the following diffusion process on $R$ [30], [31] with $u = u(p,t)$ being the diffusion variable, where $p \in R$, and $t \geq 0$ is the time variable:

$$\frac{\partial u}{\partial t} = \Delta u \quad \text{on } R \quad (7)$$

$$\frac{\partial u}{\partial n} = 0 \quad \text{on } \partial R \quad (8)$$

$$u(p,0) = v_u(p) \delta(1 - I_{0R}(p)) \quad (9)$$

where $\Delta u = u_{xx} + u_{yy}$ is the Laplacian of $u$, $\partial / \partial n$ is the oriented derivative in the direction of the exterior normal to the boundary $\partial R$, $v_u(p) = \psi(p) \cdot n$ is the magnitude of the component of the velocity normal to the boundary, $I_{0R}(p)$ is the indicator function of the set $\partial R$:

$$I_{0R}(p) = \begin{cases} 1 & \text{if } p \in \partial R \\ 0 & \text{otherwise} \end{cases}$$

and $\delta(\cdot)$ is Dirac's delta function. The diffusion process (7-9) describes, for example, the heat flow within a planar solid of shape $R$ in which the initial temperature distribution is given by (9). The Neumann boundary condition (8) specifies that there is no heat flow across the boundary $\partial R$.

Let an operator $L$ be defined by $L = -\Delta$. Using separation of variables [32], we look for a solution to (7-9) of the form $u(x,y,t) = M(x,y)N(t)$. We obtain by substituting this expression into (7, 8), the following equations:

$$LM = -\Delta M = \lambda M \quad \text{on } R \quad (10)$$

$$\frac{\partial M}{\partial n} = 0 \quad \text{on } \partial R \quad (11)$$

where $\lambda$ is the separation constant and $N' = \partial N/\partial t$. 

Authorized licensed use limited to: Univ of Calif Los Angeles. Downloaded on November 11, 2008 at 00:07 from IEEE Xplore. Restrictions apply.
Let $H$ be the Hilbert space of real square integrable functions over $R$ satisfying the boundary condition (8) with the inner product $(f, g) = \int_R f g \, ds$. It can be shown that there is an infinite number of solutions $\{M_k, \lambda_k\}_{k=1}^{\infty}$ to the eigenvalue problem (10), and that the eigenfunctions $\{M_k\}_{k=1}^{\infty}$ form a complete orthonormal system of $H$ [33]. Moreover, the eigenvalues $\lambda_k$ can be ordered in a nondecreasing sequence $0 = \lambda_1 < \lambda_2 \leq \lambda_3 \cdots$ where the general asymptotic behavior of $\lambda_k$ is given by [34] (see also [35], [36] for a review):

$$\lambda_k \sim \frac{4\pi k}{A(R)} \quad \text{as } k \to \infty,$$

(12)

which shows an asymptotic linear growth of $\lambda_k$ with $k$, depending only on the area of the region. For each specific eigenvalue $\lambda_k$, the solution to (11) becomes

$$N_k(t) = a_k \exp(-\lambda_k t),$$

(13)

where $a_k$ is a scalar. We note that the first eigenvalue of $L$ is always $\lambda_1 = 0$, with the constant function over $R$, $M_1(x,y) = C$, as its eigenfunction. By superposition, the general solution of the diffusion equation has the form [32]:

$$u(x,y,t) = \sum_{k=1}^{\infty} a_k M_k(x,y) \exp(-\lambda_k t) = \gamma + \sum_{k=1}^{\infty} a_k M_k(x,y) \exp(-\lambda_k t),$$

(14)

where $\gamma \equiv a_1 C$. The scalars $a_k$ are determined by the eigenfunction expansion of the initial condition $u(p,0)$ in the orthonormal system $\{M_k\}_{k=1}^{\infty}$. In particular,

$$a_1 = \left(\frac{u(p,0), M_1}{M_1, M_1}\right) = \frac{1}{C A(R)} \int_R u(p) \delta(1 - I(p)) \, ds = \frac{1}{A(R)} \int_R u(p) \, dl.$$

It follows form (5), (14), and (15) that

$$\lim_{t \to \infty} u(p,t) = \gamma = a_1 C = \frac{1}{A(R)} \int_R \nu_m(p) \, dl = \chi(R).$$

(16)

The precise mathematical conditions under which the above derivation holds can be found in [24], [30], [37]. We have, then, a diffusion process based solely on the component of the velocity normal to the boundary, which converges asymptotically to the object immediacy measure.

An alternative way of calculating the object immediacy measure is by integrating the normal velocity along the boundary of the object's projection on the image plane according to (5). This approach requires a perfect segmentation of the object. On the other hand, a simple variation of the diffusion technique, described in Section IV, will make it possible to circumvent this problem, while keeping the edge detection algorithm simple.

IV. IMPLEMENTATION

The results of Sections II and III suggest a scheme consisting of two different stages. In the edge detection and motion measurement stage the boundaries of the objects projected onto the image domain are defined and the velocities normal to the boundaries estimated. In the diffusion stage the output of the previous stage is used to estimate the immediacy of the objects in the scene.

We assume the image domain to be the unit square $I = [0,1]^2$, being divided by a square $n \times n$ grid. Each node of the grid is indexed by the pair $(i,j)$, with $1 \leq i, j \leq n$. We denote the image brightness at $(i,j)$ by $I(i,j)$.

A. The Edge Detection and Motion Measurement Stage

Under the assumption that the optical flow constraint equation [22] holds, the component of the motion field in the direction of the brightness gradient is given by

$$v_x(i,j) = \frac{I_y(i,j)}{\sqrt{I_x^2(i,j) + I_y^2(i,j)}},$$

(17)

where $v_x(i,j)$ is the normal velocity at $(i,j)$, and $I_x(i,j)$, $I_y(i,j)$, and $I_z(i,j)$, denote the partial derivatives of the image intensity with respect to $x$, $y$, and $z$, respectively.

To estimate the temporal derivative of the local image intensity we subtract the average grey scale values on $3 \times 3$ neighborhoods of successive frames. An approximation to the image brightness gradient direction and magnitude is determined by applying $3 \times 3$ Sobel operators [43], corresponding to the horizontal, vertical, and two diagonal axes. A moving edge with normal velocity $\nu_n(p)$ is marked at the edge detection and motion measurement stage only if the local gradient magnitude exceeds a threshold $\gamma$, and the magnitude of the estimated normal velocity is greater than $\delta$. We denote the set of all pixels marked by this procedure $E$.

B. The Diffusion Stage

Once the boundary of the object's projection on the image domain is determined, and the component of the velocity normal to the boundary is estimated, a diffusion process can be generated so as to produce, in the limit, the object immediacy measure, as described in Section III.

To realize the continuous diffusion process (7–9) we employ finite-difference methods [44]–[46]. The region $R$ is superimposed on the $n \times n$ square grid, with spacing $h = 1/n$. A time increment $\tau$ is chosen. We denote by $U(i,j,n)$ an approximation of the solution of the diffusion equation $u(x,y,t)$, evaluated at $x = ih, y = jh$, and $t = n\tau$. Let $C(i,j)$ be the subset of the 4-cell-neighborhood of $(i,j)$ such that there is no edge crossing them,

$$C(i,j) = \{(p,q) \mid |p-i| + |q-j| = 1 \text{ and } (p,q) \notin E\}. \quad (18)$$

A finite-difference approximation of the partial differential equation (7) with boundary condition (8) takes the form [45, 47]

$$U(i,j,n + 1) = \frac{1}{|C(i,j)|} \sum_{(p,q) \in C(i,j)} U(p,q,n) + n = 0, 1, \cdots.$$ 

(19)

An implicit choice of $\tau = h^2/4$ is made above. This is the maximum time increment permissible corresponding to a given...
value of $h$ such that the numerical scheme remains stable, and therefore convergent by virtue of the Lax-Richtmyer theorem [48]. The iteration method described by (19) can be interpreted as a linear network of locally connected formal neurons.

The initial condition of the diffusion network is specified by the output from the motion and edge detection stage as follows. Let a moving edge induce signals proportional to the magnitude of the component of the velocity normal to the boundary, in the immediacy of both sides of the edge, at the diffusion stage. Negative signals are induced in the direction of the edge movement, and positive ones in opposite direction. These induced signals are meant to provide the initial state of the diffusion process, which takes place in the entire image plane. If the projection of the object onto the image domain is closed as $R_I$ in Fig. 2, the positive signals correctly contribute to the computation of $\nabla \cdot R_I$, because $\mathbf{v}_x$ points outwards. The negative signals, on the other hand, will eventually fade out, as the diffusion process in an unbounded region tends asymptotically to zero. A similar description remains valid if the object’s projection is closed as $R_I$.

In reality, the image domain is finite and its boundaries may affect the results. We require zero flux across the image boundary, as described by (8). Note that an object moving into the image domain will be marked with a positive object immediacy, even when it performs a pure lateral motion with respect to the observer. It was felt that in the critical task of collision detection, under the condition of partial viewing, such false alarms can be tolerated. An additional difficulty associated with this situation is that the complement of a region $R$ is no longer unbounded. As a result, an expanding object will cause an effect of a receding background and vice versa, if the object is small compared to the image domain the effect is negligible.

To cope with the problem of incomplete boundaries the following modification was introduced. A small number $M$ of interactions are performed in the diffusion stage for each new frame. The initial condition for the current diffusion process is given by the sum of the signals induced by the velocity normal to the boundary and a constant fraction $0 \leq \alpha < 1$ of the final values attained by the diffusion process corresponding to the previous frame. This way, values at the diffusion stage become temporally integrated. While “heat” will tend to diffuse out of the region $R$ through boundary gaps, temporal integration will keep those points in the image plane which remain inside an expanding region at high “temperature.”

V. SIMULATION RESULTS

We simulated the network described in Section IV on a Sun workstation. Images are 8-bit $512 \times 512$ pixels in size. The parameters used are the same for all cases: $\alpha = 0.5$, $\gamma = 60$, $\delta = 0.9$, and $M = 100$.

The experiment (Fig. 3(a)) is described in detail in [49]. It consists of a pure translational motion of a camera towards a table on which an H-shaped bar and some pencils are placed. The focus of expansion is slightly shifted to the left from the center of the picture. Fig. 3(b) displays the result of the edge detection, while Fig. 3(c) that of the diffusion stage after four iterations (update cycles). The H-shaped bar and the two nearest pencils obtain high values of immediacy measures, represented by high (near white) gray levels. In contrast, the third pencil, which is situated at a greater distance, is not marked as an obstacle. It can be seen that the diffusion network is practically helpless where edge detection is massively incorrect (as in the midsection of the second pencil, which is merged with its background). Yet, it performs reasonably well in the face of small edge discontinuities (as the ones at the bottoms of the first and the second pencils). It can be also be seen that certain reflections from distant objects generate false immediacy signals.

Regular cameras do not use spherical projection. In principle, this could be corrected, as the origin of the distortion is purely geometric [13]. For a camera with a field of view of $25 \times 25$ degrees, such as the one used in this experiment, approximation by a spherical projection seems reasonable.

VI. CONCLUSION

We have proposed a neural mechanism for visual obstacle detection, based on a newly defined immediacy measure representing size-change information. A diffusion process on the image domain, whose initial condition is determined by the motion field normal to the object’s boundary, was shown to converge asymptotically to the immediacy measure. No differentiation of the velocity data is required. This shows the possibility of extracting useful motion information directly from the component of the velocity normal to the boundaries without prior computation of the entire motion field, as is usually done. A linear network of formal neurons was derived by means of a finite-difference approximation of the continuous diffusion process. The network was simulated and found to perform well on natural image sequences.

ACKNOWLEDGMENT

This study has been inspired by the Nap of the Earth (NoE) rotorcraft navigation problem, presented to us along with the simulation data for by Dr. Banavar Sridhar and Dr. Yair Bamiv of the Aircraft Guidance and Navigation Branch, NASA Ames Research Center.
REFERENCES


N-folded Symmetries by Complex Moments in Gabor Space and Their Application to Unsupervised Texture Segmentation

Josef Bigun and J. M. Hans du Buf

Abstract— Complex moments of the Gabor power spectrum yield estimates of the N-folded symmetry of the local image content at different frequency scales, that is, they allow to detect linear, rectangular, and circular symmetries. Manuscript received October 8, 1991; revised December 28, 1992. This work is partly supported by Thordon-CSF (France). Recommended for acceptance by Editor-in-Chief A. K. Jain.

J. Bigun and J. M. Hans du Buf are with the Signal Processing Laboratory, Swiss Federal Institute of Technology, EPFL-Ecublens, CH-1015 Lausanne, Switzerland.

IEEE Log Number 9212237.