

## Image Gradient Evolution-A Visual Cue for Collision Avoidance

Hongche Liu<sup>†‡</sup>, Tsai-Hong Hong<sup>‡</sup>, Martin Herman<sup>‡</sup>, and Rama Chellappa<sup>†</sup>  
hongche@etak.com, hongt@cme.nist.gov, herman@cme.nist.gov, chella@eng.umd.edu  
<sup>‡</sup>NIST, Intelligent Systems Division, Bldg 220, Rm B124, Gaithersburg, MD 20899  
<sup>†</sup>Center for Automation Research, University of Maryland, College Park 20742

### Abstract

*This paper is concerned with the task of visual motion-based navigation. A critical requirement of the task is the ability to estimate 3-D depth and motion from visual information. Recent studies have demonstrated that the relevant cues are contained in motion parallax or optical flow from which flow field divergence and hence time-to-contact can be extracted. We present a new concept called image gradient evolution (IGE), which utilizes the change of image spatial gradients over time as a threat cue: an approaching object induces 2-D expanding motion and causes the image spatial structure to stretch so the image gradients decrease. Based on this idea, our method offers a one-step solution directly from image gradients, instead of from optical flow and its derived properties. We use a technique that is local and linear so the implementation can be very fast. The threat map is expectedly noisy but sufficiently informative, as is seen in demonstrations on several real images. These two aspects, fast implementation and useful qualitative information, provide a viable solution to navigational tasks.*

### 1. Motivation

We start with a simple illustration to introduce the idea of the image gradient evolution (IGE) and emphasize its difference from the optical flow approach.



Fig 1.1 Approaching box

Fig 1.1 shows a diverging object in the image sequence. Our goal is to obtain an algorithm that can warn the observer of a potential collision.

Conventionally, optical flow is computed first and then the first order flow field properties (diverging or converging) are used to characterize the underlying objects' 3-D

motion. In our approach, however, the change of the image gradients over time is computed. As illustrated in Fig 2, a

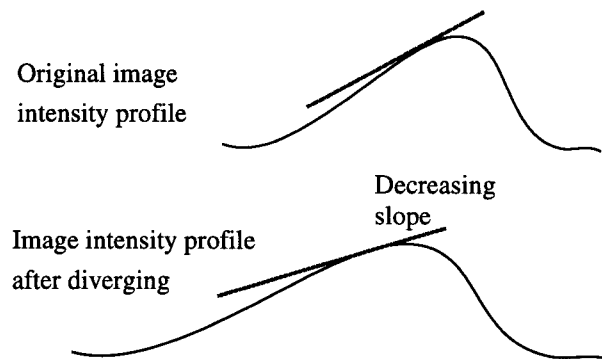


Fig 2. 1-D image gradient evolution

decreasing image intensity gradient (slope) at an image point over time signifies a diverging object. We wish to avoid processing the noisy flow data when it is evident that we can achieve the same result from IGE.

### 2. Previous Work

The looming effect is a major cue in biological vision systems to sense danger (Schiff, et al.[23]). Local motion parallax is in turn a cue for looming or divergence (Werkhoven & Koenderink[27]). Since time-to-contact is related to both divergence and 3-D scene structure, solving for time-to-contact is, in an exact sense, equivalent to recovering 3-D motion and structure. There is a plethora of literature dealing with recovering 3-D motion and structure from optical flow (Adiv[1], Bruss & Horn[4], Negahdaripour & Lee[20]), image sequences (Broida & Chellappa[3]), or features (Negahdaripour & Horn[19], Tsai & Huang[26]). It is basically an ill-posed and nonlinear problem. These methods usually use iterative optimization techniques which are often time-consuming. However, these reports have established, in theory, the feasibility of imposing extra constraints and/or using derivatives of flow to solve the problem. The intended precision of these quantitative methods is often an illusion due to the limit on ac-

curacy with which the input measures can be obtained (Thompson & Kearney[25]). Even accurate optical flow estimation is still a difficult problem.

Due to the aforementioned imprecision, there have been approaches to the avoidance of differentiating optical flow. Integration theorems such as Green's theorem (Poggio, et al.[22], Cipolla, et al.[8]), Stoke's theorem [22], and Gauss's theorem (Gupta[10]) are used to estimate first order flow parameters directly from image intensity integrals[10], image moments (and their temporal derivatives) [8], or flow integrals[22]. The integration techniques basically trade off noise sensitivity for smoothness, which arises from a single motion assumption within the integration contour. The key issue in this approach is to prevent the integration contour from going across boundaries. So additional mechanisms, probably global, are required to segment images.

Another approach models the local motion with an affine model. It uses higher order pointwise image derivatives (Nagel[18], Werkhoven & Koenderink[27]) or patchwise motion coherence (Campani & Verri [6], Bergen, et al.[2]) techniques to solve for first order motion parameters. This approach is actually aimed at accurate flow estimation and the reports make little mention of 3-D motion estimation.

Since the above two approaches do not model the 3-D structure, they do not offer sufficient information to solve for time-to-contact without additional constraints or differentiations in the general case. It has been proved that only the upper and lower bounds on time-to-contact can be derived from the above information (Subbarao[24])(Cipolla, et al.[8]). Meyer and Boutheymy[13] used temporal derivatives on the first order parameters to circumvent the problem. Essentially it is equivalent to second order derivatives, but Kalman filtering on the temporal derivatives makes the result much smoother and more practical. However, a fast implementation is relatively difficult.

Nelson & Aloimonos[21] were the first to use flow-based information for navigation. Their algorithm computes directional divergence, which is a second order flow parameter, and can be very noisy. Coombs, et al.[9] employed flow divergence for real-time obstacle avoidance. Their obstacle avoidance system currently appears to be the fastest and most reliable one using flow divergence. In their system, the time-to-contact is equivalent to divergence when carefully controlling the camera so as to approximately translate in the direction of the optical axis. The computation of divergence, however, is based on noisy quantized flow and requires temporal integration in order to interpret the result. Camus[7] implemented a real-time algorithm for time-to-contact which is quite reliable. However, the application is limited by its restrictive assumptions about motion and surfaces. Kundur & Raviv[13] pro-

posed the use of an image quality measure for the visual threat cue. Their method exploits the camera defocusing effect for navigation.

The major contribution of this paper is to pioneer the idea of IGE and its use as a cue for threat during navigation. Such a capability is embedded in a generalized 2-D motion equation that also models expansion. An integrated spatio-temporal filtering scheme is designed to estimate image derivatives in a numerically coherent manner. Using these derivatives, IGE and optical flow can quickly be estimated at the same time in a fast manner.

### 3. Generalized Motion Model

The brightness constancy equation is often interpreted in the following way:

$$\nabla I = 0 \Rightarrow I(x, y, t) = F(x-ut, y-vt). \quad (1)$$

The first step in measuring IGE is to extend the image motion model from simple 2-D translation to translation plus expansion. A 3-D point at position  $P = (X, Y, Z)$ , under perspective projection, projects to a point in the 2-D image plane,  $(x, y)$ ,

$$\begin{aligned} x &= fX/Z, \\ y &= fY/Z \end{aligned} \text{ where } f \text{ is the focal length of the projection.} \quad (2)$$

Let there be relative 3-D translational motion  $P(t) = (X + U_X t, Y + U_Y t, Z - U_Z t)$ . Hence,

$$\begin{aligned} x(t) &= f(X + U_X t)/(Z - U_Z t) \\ y(t) &= f(Y + U_Y t)/(Z - U_Z t) \end{aligned} \quad (3)$$

Brightness constancy (1) and (3) yield

$$I(x, y, t) = F\left(x\left(1 - \frac{U_Z}{Z}\right) - \frac{fU_X}{Z}, y\left(1 - \frac{U_Z}{Z}\right) - \frac{fU_Y}{Z}\right). \quad (4)$$

3-D translation only is assumed for its simplicity since 3-D rotation has no bearing on expansion (Koenderink[12]). A general motion model that includes rotation parameters can be found in (Liu, et al.[15] [16]).

To understand the generalized motion equation better, we describe equation (4) in the context of optical flow.

$$\begin{aligned} (u, v) &= \left(\frac{\partial x}{\partial t}, \frac{\partial y}{\partial t}\right) = \left(\frac{fU_X}{Z - U_Z t} + \frac{U_Z x}{Z - U_Z t}, \frac{fU_Y}{Z - U_Z t} + \frac{U_Z y}{Z - U_Z t}\right) \\ &\approx \left(\frac{fU_X}{Z} + \frac{U_Z x}{Z}, \frac{fU_Y}{Z} + \frac{U_Z y}{Z}\right) \end{aligned} \quad (5)$$

Let  $\frac{U_Z}{Z}$  be denoted by  $s$ , and  $\left(\frac{fU_X}{Z}, \frac{fU_Y}{Z}\right)$  by  $(p, q)$ . (6)

Rewriting (4) and (5):

$$I(x, y, t) = F(x(1-s) - pt, y(1-s) - qt) \quad (7)$$

$$(u, v) = (p + sx, q + sy). \quad (8)$$

Equation (8) lays out the two components of optical flow:  $(p, q)$ , and  $(sx, sy)$ . Since  $(sx, sy) = s(x, y)$ ,  $s$  is interpreted as expansion. The translation component  $(p, q)$  is

induced only by  $(U_x, U_y)$  (6), and the expansion component  $s$  only by  $U_z$  (6). When  $s = 0$ , then  $(u, v) = (p, q)$  and (7) reduces to (1).

The IGE is characterized by the following equation (refer to [15] for details),

$$\frac{\partial}{\partial x} I(x, y, t) = (1-st) \frac{\partial}{\partial x} I(x', y', 0), \begin{cases} x' = x(1-st) - pt \\ y' = y(1-st) - qt \end{cases} \quad (9)$$

An image point  $(x', y')$  at time 0 moves to  $(x, y)$  at time  $t$ . The two image gradients at the two instants are related by (9). When  $s$  is positive, it means the object is approaching ( $U_z > 0$ , in (4)), and the slope is decreasing ( $1-st < 1$ ). This coincides with our previous observation in Fig 2. In the extreme case, when  $st = 1$ ,

$$t = \frac{1}{s} = \frac{Z}{U_z} \text{ and } \frac{\partial}{\partial x} I(x, y, t) = 0. \quad (10)$$

In equation (10),  $t$  is distance  $Z$  divided by speed  $U_z$  so is interpreted as the time-to-contact. It is clear that  $s$  measures not only 2-D expansion, but also IGE and time-to-contact. We can use it as a cue for visual threat.

Note that the focus of expansion (FOE) is predefined to be at  $(0, 0)$  in (4). To complete the formulation, we modify (7) and (8) to allow the FOE to be at an arbitrary location  $(x_0, y_0)$ .

$$I(x, y, t) = F((x-x_0)(1-st) - pt + x_0, (y-y_0)(1-st) - qt + y_0) \quad (11)$$

$$(u, v) = (p + s(x-x_0), q + s(y-y_0)). \quad (12)$$

Note that equation (9) is derived based on the assumption of a parallel frontal surface, i.e., the surface normal is parallel to the optical axis. When the surface is not parallel frontal, the IGE cannot be reliably interpreted for 3-D motion. However, in order to use IGE as a qualitative cue for threat, our algorithm identifies other types of surfaces and potential boundaries as outliers. A post-smoothing stage then overwhelms the errors induced by the outliers. This technique is reasonable because ‘‘divergence due to a relatively distant object can be large, but only over a short distance in the image’’ (Nelson & Aloimonos[21]).

## 4. Algorithm and Implementation

To facilitate the estimation of IGE in the framework of the general motion model, a potent and stable image differentiation filtering scheme is needed. The set of orthogonal 3-D Hermite polynomial filters is excellent for the task.

### 4.1 Hermite polynomials

The  $n$ th Hermite polynomial  $H_n(x)$  is a solution of

$$\frac{d^2 H_n}{dx^2} - 2x \frac{dH_n}{dx} + 2nH_n = 0. \quad (13)$$

The  $H_n(x)$  are derived by Rodrigues’ formula [11]

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}. \quad (14)$$

By substituting  $G(x)$  (with variance  $\sigma^2$ ) for  $e^{-x^2}$  in (14), we generalize to Hermite polynomials with respect to the Gaussian function. Let these Hermite polynomials be denoted by  $\bar{H}_n(x)$ . Then

$$\bar{H}_n(x) = \left( \frac{1}{2^{1/2}\sigma} \right)^n H_n \left( \frac{x}{2^{1/2}\sigma} \right). \quad (15)$$

The scalar product of two functions and the  $L_2$ -norm of a function with  $G(x)$  as a weight function are defined as:

$$\langle a, b \rangle \equiv \int_{-\infty}^{\infty} G(x) a(x) b(x) dx \text{ and } \|a\| \equiv \langle a, a \rangle^{1/2}. \quad (16)$$

The orthogonality of  $\{\bar{H}_n(x)\}$  can be expressed in the following way[11]:

$$\langle \bar{H}_m, \bar{H}_n \rangle = \sigma^{-2n} n! \delta_{mn}, \quad (17)$$

The 3D case of Hermite polynomials is especially simple because they are separable.

### 4.2 Derivation of gradient constraint equations

We use the following Gaussian derivative theorem to derive motion constraint equations.

Theorem 1: A one dimensional signal  $I(x)$  can be expanded in terms of Hermite polynomials as

$$I(x) = \sum_{k=0}^{\infty} I_k \frac{\bar{H}_k(x)}{\|\bar{H}_k\|^2} \quad (18)$$

Then  $I_k = \langle I, \bar{H}_k \rangle = \langle I^{(k)}, \bar{H}_0 \rangle$ , where  $I^{(k)} = \frac{d^k I}{dx^k}$ .

Recall our motion model,

$$I(x, y, t) = F((x-x_0)(1-st) - pt + x_0, (y-y_0)(1-st) - qt + y_0) \quad (19)$$

Expand both sides with Hermite polynomials,

$$\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} I_{ijk} \frac{\bar{H}_{ijk}}{\|\bar{H}_{ijk}\|^2} = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} F_{ijk} \frac{\bar{H}_{ijk}}{\|\bar{H}_{ijk}\|^2} \text{ then} \\ I_{ijk} = \langle I, \bar{H}_{ijk} \rangle = F_{ijk} = \langle F, \bar{H}_{ijk} \rangle \quad (20)$$

Equating  $I_{ij1}$  to  $F_{ij1}$  and using Theorem 1, we derive

$$I_{ij1} \approx -u I_{(i+1)j0} - v I_{i(j+1)0} - (i+j)s I_{ij0} \text{ where } (u, v) \text{ are defined in (12)}. \quad (21)$$

The key concept in the derivation is the use of the recursive relation of the Hermite polynomial (13) to obtain a lin-

ear equation of image derivatives, optical flow ( $u, v$ ) and IGE  $s$  [15]. Using 00, 01, 10 for  $ij$ , we can derive three equations up to the second order and solve the linear system. Notice that both optical flow and the IGE can be computed directly from the image derivatives. In our implementation, we use six equations up to the third order to form a least square formulation. The reason is that the residual is an excellent reliability measure. In fact, if we consider the residual as the extent to which the motion model is violated, it can be used to indicate noise, non-frontal surfaces, and boundaries (Liu, et al.[17]). Since  $s$  is noisy, our implementation exploits smoothing with confidence weighting, i.e., extra steps of smoothing on  $s$  with the reciprocal of the residual as weighting. This smooths out noise but prevents smoothing across boundaries. It also overwhelms the errors due to non-frontal surfaces as long as there is a portion within the object that is parallel frontal or nearly parallel frontal.

## 5. Experiments

The following figures (Fig 3.-Fig 5.) show one image of the sequence and its 3-D perspective threat map based on IGE  $s$ . In the 3-D threat maps, elevation is used to depict threat. The elevated area represents closer objects. The threat value is thresholded to enhance 3-D structure perception. For example, in Fig 3, the cereal box stands out in the

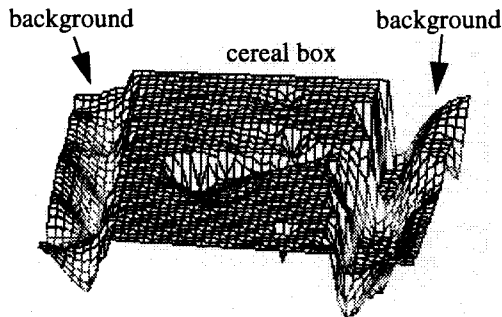


Fig 3. Threat map of the approaching box

center because most of the threat values within the box are above the threshold; in Fig 4, the threat map is gradually elevated from the Yosemite valley to the mountain in the lower left corner. The sky is perceived as mostly safe except at some areas where the clouds change brightness irregularly and deceive the algorithm. A navigation algorithm can use this threat map to avoid heading towards the lower left corner; in Fig 5, the metal plate in the lower left corner is extracted; the Coke can and the platform in the center are also partially visible; the pole on the right side is visible at both of its ends. The fact that the metal plate is detected and the hole remains intact demonstrates the effect

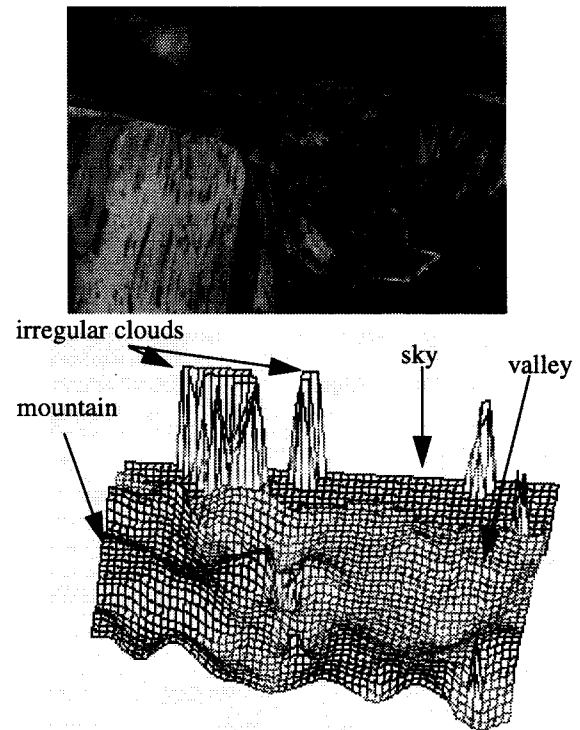


Fig 4. Yosemite and threat map

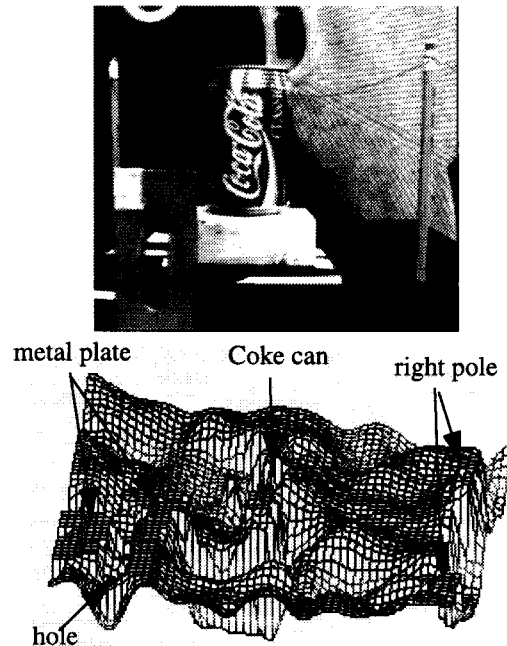


Fig 5. NASA sequence and threat map

of smoothing with confidence weighting. We are currently working on a real-time implementation of the algorithm. On 64x64 images, IGE without smoothing can be expected to run at 3-4 frames per second on a HyperSparc 10 MP board. The amount of confidence smoothing is dependent

on the image noise and may take a little more time than required for computing the IGE. Although the threat map is noisy and currently its resolution is limited by noise and irregular brightness changes, it already provides useful information for navigation.

## 6. Conclusion

The use of the IGE eliminates the need to process noisy optical flow. Image gradient evolution has been shown to be a useful cue for threat. Our algorithm builds a dense qualitative threat map based on IGE. For navigation, we would rather compute useful information quickly than exact information slowly. That is what our algorithm achieves.

## Reference

- [1] Adiv, G., "Determining Three-Dimensional Motion and Structure from Optical Flow Generated by Several Moving Objects", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 7, no. 4, pp. 384-401, 1985.
- [2] Bergen, J.R., Burt, P.J., Hingorani, R., Peleg, S., "A Three-Frame Algorithm for Estimating Two-Component Image Motion", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 9, pp. 886-896, 1992.
- [3] Broida, T.J., Chellappa, R., "Estimation of Object Motion Parameters from Noisy Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 1, pp. 90-99, 1986.
- [4] Bruss A.R., Horn, B.K., "Passive Navigation", *Computer Vision, Graphics, and Image Processing: Image Understanding*, vol. 21, no. 1, pp. 3-20, July 1983.
- [5] Burlina, P., Chellappa, R., "Time-to-X: Analysis of Motion Through Temporal Parameters", *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 461-468, Seattle, WA, 1994.
- [6] Campani, M. and Verri, A., "Motion Analysis from First-Order Properties of Optical Flow", *Computer Vision, Graphics, and Image Processing: Image Understanding*, vol. 50, no. 1, pp. 90-107, July 1992.
- [7] Camus, T., "Calculating Time-to-Contact Using Real-Time Quantized Optical Flow", NISTIR-5609, April, 1995.
- [8] Cipolla, R., Blake, A., "Surface Orientation and Time to Contact from Image Divergence and Deformation", *Proceedings of IEEE European Conference on Computer Vision, Italy*, pp. 187-202, May, 1992.
- [9] Coombs, D., Herman M., Hong T., Nashman, M., "Real-time Obstacle Avoidance Using Central Flow Divergence and Peripheral Flow", *Proceedings of IEEE International Conference on Computer Vision, Cambridge, MA, June, 1995*.
- [10] Gupta, N., "Recovering Shape and Motion from a Sequence of Images", PH.D. Dissertation, University of Maryland, April, 1993.
- [11] Hashimoto M, and Sklansky, J. "Multiple-Order Derivatives for Detecting Local Image Characteristics", *Computer Vision, Graphics, and Image Processing*, vol. 39, pp. 28-55, 1987.
- [12] Koenderink, J.J., "Optic Flow", *Vision Research*, vol. 26, pp. 161-180, 1986.
- [13] Kundur, S.R., Raviv, D., "An Image-Based Texture-Independent Visual Motion Cue for Autonomous Navigation", NISTIR 5567, April 1995; *Proceedings of IEEE International Symposium on Computer Vision, Coral Gable*, pp. 527-532, Nov 1995.
- [14] Meyer, F., Boutheimy, P., "Estimation of Time-to-Collision Maps from First Order Motion Model and Normal Flows", *Proceedings of IEEE International Conference on Pattern Recognition*, pp. 78-82, 1992.
- [15] Liu, H., Hong, T., Herman, M. and Chellappa, R., "A Generalized Motion Model and Spatio-Temporal Filters for 3-D Motion Interpretations", PH.D. thesis, University of Maryland, 1995.
- [16] Liu, H., Hong, T., Herman, M. and Chellappa, R., "A General Motion Model and Spatio-temporal Filters for Computing Optical Flow", University of Maryland-TR -3365, November 1994; to appear in *International Journal of Computer Vision*.
- [17] Liu, H., Hong, T., Herman, M. and Chellappa, R., "Motion-Model-Based Boundary Extraction", University of Maryland TR-3414, February, 1995; *Proceedings of IEEE International Symposium on Computer Vision, Coral Gable*, pp. 587-592, Nov 1995.
- [18] Nagel, H. H., "Direct Estimation of Optical Flow and of Its Derivatives", *Artificial and Biological Vision Systems*, Ed. Orban, G.A., Nagel, H. H., pp. 193-224, 1992.
- [19] Negahdaripour, S. Horn, B.K., "Direct Passive Navigation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 1, pp. 168-176, 1987.
- [20] Negahdaripour, S. Lee, S., "Motion Recovery from Image Sequences using First-Order Optical Flow Information", *Proceedings of IEEE Workshop on Visual Motion, Princeton*, pp. 132-139, Oct, 1991.
- [21] Nelson, R. and Aloimonos, J., "Obstacle Avoidance Using Flow Field Divergence", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 10, pp. 1102-1106, 1989.
- [22] Poggio, T., Verri, A., Torre, V., "Green Theorems and Qualitative Properties of the Optical Flow", AI Memo 1289, MIT AI Laboratory, April, 1991.
- [23] Schiff, W., Caviness, J.A., Gibson, J.J., "Persistent Fear Response in Rhesus Monkeys in Response to the Optical Stimulus of Looming", *Science*, vol. 136, pp. 982-983, 1962.
- [24] Subbarao, M., "Bounds on Time-to-Collision and Rotational Component from First Order Derivatives of Image Flow", *Computer Vision, Graphics, and Image Processing: Image Understanding*, vol. 50, no. 1, pp. 329-341, 1990.
- [25] Thompson, W.B., Kearney, J.K., "Inexact Vision", *Proceedings of the Workshop on Motion, Representation, and Analysis*, pp.15-21, May 1986.
- [26] Tsai, R.Y., Huang, T.S., Zhu, W.L., "Estimating Three-Dimensional Motion Parameters of a Rigid Planar Patch II: Singular Value Decomposition", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 30, Aug 1982.
- [27] Werkhoven, P., Koenderink, J.J., "Extraction of Motion Parallax Structure in the Visual System I", *Biological Cybernetics*, vol. 63, pp. 185-191, 1990.