NISTIR 4808

A QUANTITATIVE APPROACH TO LOOMING

Daniel Raviv

Florida Atlantic University
The Robotics Center and
The Electrical Engineering Department
Boca Raton, Florida
and
Sensory Intelligent Group
Robot Systems Division

U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards
and Technology
Robot Systems Division
Bldg. 220 Rm. B124
Gaithersburg, MD 20899

NIST

NISTIR 4808

A QUANTITATIVE APPROACH TO LOOMING

Daniel Raviv

Florida Atlantic University
The Robotics Center and
The Electrical Engineering Department
Boca Raton, Florida
and
Sensory Intelligent Group
Robot Systems Division

U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards
and Technology
Robot Systems Division
Bldg. 220 Rm. B124
Gaithersburg, MD 20899

April 1992

U.S. DEPARTMENT OF COMMERCE
Barbara Hackman Franklin, Secretary

TECHNOLOGY ADMINISTRATION
Robert M. White, Under Secretary for Technology

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
John W. Lyons, Director

ABSTRACT

The visual looming effect, i.e., expansion of object's size in the retina, has been shown to be very important when action is taking place. Most existing research work on "looming" has been done by psychologists and is qualitative or limited-quantitative. In order to use the existing results for robotics applications, a more mathematically-oriented approach for looming is needed. In this paper we take a quantitative approach to understanding looming. First we define looming mathematically and show geometrical properties of objects that produce the same value of looming. Then we explain how to measure looming in the general case of motion using optical flow and the change of the projection of a 3-D object on the retina. It is shown how a multiresolution logarithmic retina simplifies the measurement of looming, and how the results can be combined with previous work on visual fields. We suggest a new representation of space based on looming and the so called Equal Flow Circles (EFCs).

This research was supported in part by a grant from the National Science Foundation to Florida Atlantic University (Grant #IRI-9115939).

1. INTRODUCTION

The Random House dictionary defines *looming* as: "Coming into view in indistinct and enlarged form; rising before the vision with an appearance of great or portentous size". Although this definition is a qualitative one, it is clear that it deals with expansions of objects on the retina. Usually, expansion of object's size on the retina is a result of a decreasing distance over some period of time between the observer and the object.

The perception of looming which is an indication for possible collision is critical to survival for creatures of nature. It is necessary for locomotion in a complex natural 3-D world. The reaction to this visual stimulus is the result of some kind of "perceived threat", i.e., the measured relative rate of expansion of objects on the retina corresponds to a visual timing parameter that causes the subject to defensively react to reduce the visual threat.

Looming and looming-related actions have been studied by many researchers, mainly psychologists [1-24], but most of the work is qualitative or limited-quantitative [9]. The looming effect [20,45], which is the result of optical expansion of objects, has shown to cause defensive reaction in several animals as well as babies [6,20]. In several reported experiments the subjects' use of visual information in the last few hundreds of milliseconds before a collision occurs was examined.

In order to use the existing results for robotics applications, a more mathematically-oriented approach for looming is needed.

Lee [9,10] and his colleagues [11-14] showed how to quantitatively measure the "time-to-contact" from optical flow for an observer who undergoes only translation when the optical axis coincides with the motion vector. Extensions of the time-to-contact concept were recently reported in [25,26]. Our definition of looming is related to, but different from, the time-to-contact. The time-to-contact according to [9] is the time it takes for an observer to hit a specific *plane* perpendicular to the direction of motion, i.e., it deals with the relative change of the *depth* (as opposed to *range).*

Assume an observer moving in a 3-D environment filled with several balls (Figure 1). During the motion of the observer relative to the balls, both the size of each ball and the location its centroid, when projected into the image, are continuously changing. If the distance between a ball and the center of the camera decreases then the projected image of the ball will increase in area. The relative rate of expansion of the area over time of the imaged ball causes the "looming" effect and it is proportional to the time derivative of the relative distance (range) between the observer and the ball *divided* by the relative distance (range) itself. The looming is a measurable variable and can be extracted *directly* from a sequence of 2-D images using optical flow, relative change in area, etc.

In this paper we approach looming from quantitative as well as qualitative points of view. We start with a mathematical definition of looming, followed by an expression for looming in vector form. We show that looming is rotation independent and that there are equal looming spheres each of which corresponds to a value of looming. In other words, each point on a particular equal-looming sphere appears as having the same looming regardless of the instantaneous rotation of the camera. We introduce several methods for measuring looming, using optical flow as well as the expansion of the projection of objects on the retina. We describe camera-centered coordinate systems that are used here, and derive expressions that relate optical flow and distance of a point to six-degree-of-freedom camera motion. When the optical flow constraint (as described by Horn and Schunck [36]) is added to these equations, a pixel-based closed-form solution for the looming of a point in the camera coordinate system is obtained. The rate of expansion of projected objects is another method introduced to measure looming.

It is shown that by using a logarithmic retina to measure looming, many computations become significantly simpler, and that the computations are independent of the object's size. The analysis becomes even simpler when dealing with looming and fixation. A mapping of space is suggested using equal looming spheres and Equal Flow Circles (EFCs, as described in [63]).

The work described in this paper is highly related to the area called "Active Vision", in particular purposive quantitative as well as qualitative vision, since it can be used to actively control the viewer's parameters. [30,31,33-35,57,66].

2. MATHEMATICAL DEFINITION OF LOOMING

Let a camera and an object (e.g., a ball) move arbitrarily in a 3-D environment (Figure 2a). Then at two different time instants they may be in different relative locations. The distance (range) from the center of the ball to the pinhole point of the camera may change, thus resulting in different sizes of projected images. Figure 2b shows a 3-D ball and its images at two different time instants. The looming effect is caused by the expansion of the projected ball. Intuitively, the relative rate of expansion, i.e., the relative change of the projection of the ball $\frac{A_2-A_1}{A_1}$ over the period of time t_2-t_1 in which the change takes place, is highly related to the relative change of the range $\frac{R_1-R_2}{R_1}$ during the same period of time. (In fact, for objects that are small relative to their range from the observer, they are proportional, i.e., differ by a scale factor.) Thus one can define looming in terms of the relative change in range instead of relative change in the object's projection.

In order to mathematically define looming and suggest ways to measure it we shall consider infinitesimally small 3-D balls. Let the relative distance between the observer and a point P (the center of the infinitesimally small ball) at time instant t_1 be R_1 and at time instant t_2 be R_2 (Figure 3). Then, we define looming L as the *negative* value of the time derivative of the relative distance between the observer and the point P, divided by the relative distance R, i.e.,

$$L = -\lim_{\Delta t \to 0} \frac{\frac{(R_2-R_1)}{\Delta t}}{R_1} \tag{1}$$

where $\Delta t = t_2 - t_1$. Or

$$L = -\frac{\dot{R}}{R} \tag{2}$$

where dot denotes derivative with respect to time. The reason for the *negative sign* in Equations (1) and (2) is to associate image *expansion* with *positive* looming.

This definition allows the use of the term "looming of a point" for describing the value of L (i.e., $\frac{-\dot{R}}{R}$) of a point.

3. LOOMING IN VECTOR FORM

3.1 STATIONARY ENVIRONMENT

Assume a moving observer in a stationary environment. Let the instantaneous vector from the observer to a stationary point P be **R**, the instantaneous translational velocity vector of the camera be **t** and the instantaneous angular velocity vector of the camera be $\boldsymbol{\omega}$ (Figure 4). Then the expression for the velocity vector **V** of point P in camera coordinates is (see for example [51]):

$$\mathbf{V} = -\mathbf{t} - \boldsymbol{\omega} \times \mathbf{R} \tag{3}$$

The general expression for the looming L in vector form is:

$$L = -\frac{\mathbf{V}\cdot\mathbf{R}}{\mathbf{R}\cdot\mathbf{R}} \tag{4}$$

or

$$L = -\frac{\mathbf{V}\cdot\mathbf{R}}{|R|^2}$$

The last expression is *identical* to the expression in Equation (2) for the following reasons: $\frac{\mathbf{R}}{|R|}$ is a unit vector along the line that connects the observer with the point in 3-D. $\mathbf{V}\cdot\frac{\mathbf{R}}{|R|}$ is the projection of the velocity vector **V** along the $\frac{\mathbf{R}}{|R|}$ direction, i.e., it is $\dot{R}$. Hence $\frac{\mathbf{V}\cdot\mathbf{R}}{|R|^2} = \frac{\dot{R}}{R}$.

By substituting (3) into (4) the expression for the looming L becomes:

$$L = -\frac{(-\mathbf{t} - \boldsymbol{\omega} \times \mathbf{R})\cdot\mathbf{R}}{\mathbf{R}\cdot\mathbf{R}} \tag{5}$$

$(\boldsymbol{\omega} \times \mathbf{R})\cdot\mathbf{R} = 0$ (since the vector $\boldsymbol{\omega} \times \mathbf{R}$ is always perpendicular to the vector **R**) and

thus expression (5) becomes:

$$L = \frac{\mathbf{t} \cdot \mathbf{R}}{\mathbf{R} \cdot \mathbf{R}} \tag{6}$$

In the expression for looming (Equation (6)) neither the distance nor the velocity between the observer and the point in space need to be explicitly known. Only the *ratio* between the two dot products needs to be known. The expression is time dependent only, i.e., L is measured in [$time^{-1}$] units.

Note that the result for L in Equation (6) is a scalar which is dependent on the camera translational component of the velocity vector, but *independent* of any camera rotation.

3.2 MOTION OF A CAMERA AND POINTS

In the case where both the camera and the point are moving, we can always analyze it as if the point is stationary and the camera undergoes translation and rotation. The translation vector of the camera relative to the "artificially" stationary point is clearly the result of subtracting the point translation vector from the camera translation vector. For looming computation the rotation is not relevant, and thus only the *relative* translational information between the observer and the point is relevant for any kind of relative motion between the observer and the point. The same result is true for motion of a camera and a 3-D rigid object.

4. EQUAL LOOMING SURFACES

In this section we answer the following question: "What is the location of points in 3-D space that result in the same looming value L for any motion of the camera?"

4.1 GENERAL MOTION OF CAMERA

Let us set the value of the looming expression (Equation (6)) to a constant and obtain the equal looming surfaces. If the angle between $\mathbf{t}$ and $\mathbf{R}$ is α then Equation (6) becomes:

$$L = \frac{|\mathbf{t}|}{R}\cos\alpha \tag{7}$$

where $|\mathbf{t}|$ is the absolute value of the vector $\mathbf{t}$. Without loss of generality attach an

XYZ coordinate system to the observer such that the Y axis coincides with the instantaneous vector $\mathbf{t}$ as in Figure 5. In this case $\cos\alpha = \frac{y}{\sqrt{x^2+y^2+z^2}}$ and $R = \sqrt{x^2+y^2+z^2}$. Substituting these expressions into Equation (7) yields:

$$L = \frac{|t|\, y}{x^2+y^2+z^2} \tag{8}$$

or

$$x^2 + (y - \frac{|t|}{2L})^2 + z^2 = (\frac{|t|}{2L})^2 \tag{9}$$

Equation (9) is an equation of a sphere whose center is located at $(0,\frac{|t|}{2L},0)$ and its radius is $\frac{|t|}{2|L|}$. The center of the sphere lies on the instantaneous translation vector, and the observer lies on a point of this sphere (the origin of the coordinate system in Figure 5). Note that this sphere corresponds to a particular looming value, i.e., all points on this sphere result in the same value of L. Depending on $|t|$ this sphere will expand or shrink in 3-D. Figure 6 shows points in 3-D that produce the same looming value, and Figure 7 shows the location of the equal looming points for different looming values. Each circle in Figures 6 and 7 is a section of an equal looming sphere.

4.2 POSITIVE-, NEGATIVE-, AND ZERO-LOOMING SURFACES

Of the equal looming surfaces some correspond to positive values of looming, others correspond to negative values of looming and there is a plane (i.e., a sphere with infinite diameter) that corresponds to zero looming. Figure 8 shows the three types of surfaces. Points on the plane which is perpendicular to the instantaneous translational vector and passes through the pinhole point of the camera produce zero looming. Points on the hemisphere in front of the moving camera (i.e., in front of the zero-looming plane) produce positive values of looming, and points on the hemisphere behind the moving camera (i.e., behind the zero-looming plane) produce negative looming values. Note that the smaller the sphere the higher the absolute value of the looming. The different points shown in Figure 8 correspond to different L which depends on their location. Note that at $\theta = 45°$ the looming value is maximum and at $\theta = -45°$

it is minimum. At these two angles the size of the sphere is minimal. The looming as a function of the angle θ (as defined in Figure 8) is graphically described in Figure 9.

5. LOOMING AND SCALED SPACE

As has been shown previously, the radii of the equal looming spheres are functions of time (measured in reverse unit of time) and are *not* measured in length units. This means that different spheres in 3-D may result in the same looming value when dealing with different instantaneous translational velocity. For some direction θ, a far point with large value of $|t|$ may have the same value of L as a close point with small value of $|t|$. In the scaled space, i.e., in the time domain where the coordinates are $\frac{X}{|t|}, \frac{Y}{|t|}$ and $\frac{Z}{|t|}$ (time-based coordinates) a single sphere represents a single looming value regardless of the exact translational velocity.

6. MEASUREMENT OF LOOMING

We show two basic ways for measuring looming. One is based on optical flow where the derivation is for any six-degree-of-freedom motion of a camera in a stationary environment. The second way is based on the relative rate of expansion of the objects' projections and holds for any kind of relative motion between the camera and the object.

6.1 USING OPTICAL FLOW

6.1.1 EQUATIONS OF MOTION AND OPTICAL FLOW

This section describes the equations that relate a point in 3-D space to the projection of that point in the image for general six-degree-of-freedom motion of the camera. The following derivation is for motion in a stationary environment. Some of the equations are well known [51,52].

Assume a moving camera in a stationary environment with the coordinate system fixed with respect to the camera as shown in Figure 10. Assume a pinhole camera model, such that the pinhole of the camera is at the origin of the coordinate system. We derive the optical flow components in a spherical coordinate system ($R\,\theta\phi$). In this frame, angular velocities ($\dot{\theta}$ and $\dot{\phi}$) of any point in space, say P, are identical to the

optical flow values $\dot{\theta}$ and $\dot{\phi}$ of the projection of P (i.e., P') in the image domain (Figure 11).

The relationship between the image flow values in the spherical domain is given by the following equations. Let the instantaneous cartesian coordinates of the point *P* be $\mathbf{R} = (X,Y,Z)^T$ (where the superscript *T* denotes transpose). Let the instantaneous translational velocity of the camera be $\mathbf{t} = (U,V,W)^T$ and the instantaneous angular velocity be $\boldsymbol{\omega} = (A,B,C)^T$. Then the velocity vector **V** of the point *P* with respect to the *XYZ* coordinate system is as shown in Equation (3), i.e.,

$$\mathbf{V} = -\mathbf{t} - \boldsymbol{\omega} \times \mathbf{R}$$

or:

$$V_X = -U - BZ + CY \tag{10}$$

$$V_Y = -V - CX + AZ \tag{11}$$

$$V_Z = -W - AY + BX \tag{12}$$

where V_X, V_Y, and V_Z are the components of the velocity vector **V** along the *X*, *Y*, and *Z* directions respectively. Let $s_\theta = \sin\theta$, $c_\theta = \cos\theta$, $s_\phi = \sin\phi$, $c_\phi = \cos\phi$, and $R = |\mathbf{R}|$. To convert from *XYZ* to $R\theta\phi$ coordinates we use the relations:

$$X = R\, c_\phi\, c_\theta \tag{13}$$

$$Y = R\, c_\phi\, s_\theta \tag{14}$$

$$Z = R\, s_\phi \tag{15}$$

In order to find the optical flow of a 3-D point in $R\theta\phi$ coordinates, we use the following relations and transformations (see [52] and Figure 10). Let V_R, V_θ, and V_ϕ be the components of the vector **V** in spherical coordinates, then

$$\begin{bmatrix} V_R \\ V_\theta \\ V_\phi \end{bmatrix} = \begin{bmatrix} c_\phi & 0 & s_\phi \\ 0 & 1 & 0 \\ -s_\phi & 0 & c_\phi \end{bmatrix} \begin{bmatrix} c_\theta & s_\theta & 0 \\ -s_\theta & c_\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} V_X \\ V_Y \\ V_Z \end{bmatrix} \tag{16}$$

Also from [52]

$$V_R = \dot{R}, \tag{17}$$

$$V_\theta = R\dot{\theta}\, c_\phi, \tag{18}$$

$$V_\phi = R\dot{\phi}, \tag{19}$$

where dot denotes derivative with respect to time.

Using equations (10)-(19) the following expressions are obtained:

$$\begin{bmatrix} \dot{R} \\ R\ \dot{\theta}\ c_\phi \\ R\ \dot{\phi} \end{bmatrix} = \begin{bmatrix} c_\phi c_\theta & c_\phi s_\theta & s_\phi \\ -s_\theta & c_\theta & 0 \\ -s_\phi\, c_\theta & -s_\phi\, s_\theta & c_\phi \end{bmatrix} \begin{bmatrix} -U - B\ R\ s_\phi + C\ R\ c_\phi\, s_\theta \\ -V - C\ R\ c_\phi\, c_\theta + A\ R\ s_\phi \\ -W - A\ R\ c_\phi\, s_\theta + B\ R\ c_\phi\, c_\theta \end{bmatrix} \tag{20}$$

There are four unknowns in Equation set (20): R,$\dot{R}$, $\dot{\theta}$, and $\dot{\phi}$. For each pixel, θ and ϕ are known. The motion parameters (A,B,C,U,V,W) are also known.

6.1.2 THE OPTICAL FLOW CONSTRAINT

If the brightness I varies smoothly with θ, ϕ, and t (time), then we obtain the optical flow constraint equation using the chain rule (see [36]):

$$\frac{\partial I}{\partial \theta}\dot{\theta} + \frac{\partial I}{\partial \phi}\dot{\phi} + \frac{\partial I}{\partial t} = 0 \tag{21}$$

or

$$I_\theta\,\dot{\theta} + I_\phi\,\dot{\phi} + I_t = 0 \tag{22}$$

where

$$I_\theta = \frac{\partial I}{\partial \theta} \qquad I_\phi = \frac{\partial I}{\partial \phi} \qquad I_t = \frac{\partial I}{\partial t}.$$

When dealing with digital images I_θ, I_ϕ, and I_t can be approximated from the image sequence as follows. For a given pixel (θ_i,ϕ_j) in the image at time instant t_k

$$\frac{\partial I}{\partial \theta_i} \approx \frac{I(\theta_i,\phi_j,t_k) - I(\theta_{i-1},\phi_j,t_k)}{\theta_i - \theta_{i-1}} \tag{23}$$

$$\frac{\partial I}{\partial \phi_j} \approx \frac{I(\theta_i,\phi_j,t_k) - I(\theta_i,\phi_{j-1},t_k)}{\phi_j - \phi_{j-1}} \tag{24}$$

$$\frac{\partial I}{\partial t_k} \approx \frac{I(\theta_i,\phi_j,t_k) - I(\theta_i,\phi_j,t_{k-1})}{t_k - t_{k-1}} \tag{25}$$

6.1.3 LOOMING COMPUTATIONS

Equation set (20) together with equation (22) form a set of four equations with four unknowns R, $\dot{R}$, $\dot{\theta}$, and $\dot{\phi}$. Solving these equations for R yields:

$$R = \frac{(Us_\theta - Vc_\theta)I_\theta + c_\phi(Us_\phi c_\theta + Vs_\phi s_\theta - Wc_\phi)I_\phi}{-c_\phi I_t - c_\phi(Bc_\theta - As_\theta)I_\phi + (Cc_\phi - s_\phi(Bs_\theta + Ac_\theta))I_\theta} \tag{26}$$

Solving for $\dot{R}$ yields:

$$\dot{R} = -Uc_\phi c_\theta - Vc_\phi s_\theta - Ws_\phi \tag{27}$$

And by combining Equations (26) and (27) we get:

$$L = -\frac{\dot{R}}{R} = -\frac{(-Uc_\phi c_\theta - Vc_\phi s_\theta - Ws_\phi)(-c_\phi I_t - c_\phi(Bc_\theta - As_\theta)I_\phi + (Cc_\phi - s_\phi(Bs_\theta + Ac_\theta))I_\theta)}{(Us_\theta - Vc_\theta)I_\theta + c_\phi(Us_\phi c_\theta + Vs_\phi s_\theta - Wc_\phi)I_\phi} \tag{28}$$

If the translation vector t forms $\theta=\theta_t$ and $\phi=\phi_t$ angles in the $R-\Theta-\Phi$ coordinates (Figure 10) then U, V, and W can be expressed as:

$$U = |t|c_{\phi_t}c_{\theta_t} \tag{29}$$

$$V = |t|c_{\phi_t}s_{\theta_t} \tag{30}$$

$$W = |t|s_{\phi_t} \tag{31}$$

Using the last three expressions, Equation (28) can be rewritten as:

$$L = \frac{(c_{\phi_t}c_{\theta_t}c_\phi c_\theta + c_{\phi_t}s_{\theta_t}c_\phi s_\theta + s_{\phi_t}s_\phi)(-c_\phi I_t - c_\phi(Bc_\theta - As_\theta)I_\phi + (Cc_\phi - s_\phi(Bs_\theta + Ac_\theta))I_\theta)}{(c_{\phi_t}c_{\theta_t}s_\theta - c_{\phi_t}s_{\theta_t}c_\theta)I_\theta + c_\phi(c_{\phi_t}c_{\theta_t}s_\phi c_\theta + c_{\phi_t}s_{\theta_t}s_\phi s_\theta - s_{\phi_t}c_\phi)I_\phi} \tag{32}$$

The meaning of solution (32) is that if the location of the pixel in the image, the instantaneous rotational motion parameters of the camera, the instantaneous translational *direction* of motion, and the spatial and temporal intensity-changes (i.e., changes in brightness) are given, then the looming of the corresponding point in 3-D can be obtained.

At first inspection, 33 multiplications and divisions, and 11 summations and subtractions appear to be necessary for each pixel at each instant of time. However, for a given pixel, θ and ϕ are constants and therefore all functions of θ and ϕ such as $c_\theta c_\phi$ can be precomputed. Also, the rate of change of the motion parameters A, B, C and ϕ_t, θ_t is usually much lower than the rate of change of the image variables I_θ, I_ϕ and I_t. Thus the update rate of the motion parameters can be significantly smaller than the computation of L. For example, for a translation in the XY plane, and rotation about the Z axis (i.e., A, B, and W are zero), U, V, and C may be constants for some period of time but I_θ, I_ϕ and I_t change. Practically, accelerations (either translational or rotational) are limited due to mass (or inertia) constraints. Thus, any function of $\theta, \phi, \phi_t, \theta_t, A, B, C$ will typically have an update rate that is an order of magnitude smaller than the rate of change of the pixel intensity (brightness).

Equation (32) can be rewritten as follows:

$$L = \frac{a_0(a_3 I_t + a_4 I_\phi + a_5 I_\theta)}{a_1 I_\theta + a_2 I_\phi} \tag{33}$$

where

$$a_0 = c_{\phi_t} c_{\theta_t} c_\phi c_\theta + c_{\phi_t} s_{\theta_t} c_\phi s_\theta + s_{\phi_t} s_\theta \tag{34}$$

$$a_1 = c_{\phi_t} c_{\theta_t} s_\theta - c_{\phi_t} s_{\theta_t} c_\theta \tag{35}$$

$$a_2 = c_\phi (c_{\phi_t} c_{\theta_t} s_\phi c_\theta + c_{\phi_t} s_{\theta_t} s_\phi s_\theta - s_{\phi_t} c_\phi) \tag{36}$$

$$a_3 = - c_\phi \tag{37}$$

$$a_4 = c_\phi (A s_\theta - B c_\theta) \tag{38}$$

$$a_5 = C c_\phi - s_\phi (B s_\theta + A c_\theta) \, . \tag{39}$$

By observing equation (33) it is clear that for given I_θ, I_ϕ and I_t, at the fast sampling rate there are only six multiplications, one division, and three summations per pixel. (Note the the number of multiplication can be reduced to five by multiplying a_0 by a_3, a_4 and a_5.)

For each point in space there is a set of four equations that result in a looming value of a point in space. The expression for L that corresponds to a pixel in the image can be computed using a pixel-based processor, and special purpose computers may run in parallel to compute looming values of all visible points in space. For each pixel there may be a "fast" computer, and a "slow" computer. The "fast computer" extracts I_θ, I_ϕ, and I_t, and computes the looming L that corresponds to this pixel. The "slow" computer generates a_0,a_1,a_2,a_3,a_4 and a_5 from Equations (34)-(39) at a low update rate. A suggestion for a structure of such a "multi-rate" computer can be found in [64].

The problem with this approach is that the optical flow is highly dependent on both the translation and rotation of the camera, and so the expression for the looming L in Equation (32) contains rotation parameters.

For the purpose of analyzing looming this approach is quite limited. Practically there are other severe limitations which are related to the measurements of spatial and temporal gradients.

6.1.4 SPECIAL CASE 1

In order to intuitively understand expression (32), we show a special case. Suppose that the camera undergoes translation U and V in the instantaneous XY plane and rotation C about the Z axis only, i.e., $A = B = W = 0$. Assume $\phi = 0$, i.e., the pixel which is analyzed lies in the XY plane. Substituting these values in Equation (32) results in the following expression for L:

$$L = \frac{(c_{\theta_t}c_\theta + s_{\theta_t}s_\theta)(-I_t + CI_\theta)}{(c_{\theta_t}s_\theta - s_{\theta_t}c_\theta)I_\theta} \tag{40}$$

or (assuming $I_\theta \neq 0$)

$$L = \frac{(c_{\theta_t}c_\theta + s_{\theta_t}s_\theta)(-\dot{\theta} + C)}{(c_{\theta_t}s_\theta - s_{\theta_t}c_\theta)} \tag{41}$$

where $\dot{\theta} = \frac{I_t}{I_\theta}$. Given that $\dot{\theta} \neq C$, Equation (41) is a solution for any point in the XY plane when translating in this plane and rotating about the Z axis.

6.1.5 SPECIAL CASE 2

For the case where the camera undergoes translation only along the optical axis, the expression for looming for points in the $X-Y$ plane becomes very simple, i.e.,

$$L = -\frac{\dot{\theta}}{cot\,\theta} \tag{42}$$

An alternative method for computing the optical flow is described in [53,64].

6.2. DIRECT CALCULATIONS OF LOOMING (FROM RELATIVE RATE OF EXPANSION)

This approach deals with any relative motion between the camera and the object. It is based on measuring the looming at the image region (or object) level (rather than the pixel level). The relative rate of expansion of an object in the image is proportional to the looming as defined in section 2.

6.2.1 THE 3-D BALL SCENARIO

The concept of measuring looming from the relative rate of expansion can be explained using a 3-D ball example. Figure 12 shows an observer moving relative to a 3-D ball. As a result, the projection of the ball on the retina may be different in size at two different time instants t_1 and t_2. Let the relative speed along the ball-observer direction be $|t|$, the radius of the ball r, and the distance from the center of the ball to the observer R. At $t=t_1$

$$\sin\beta = \frac{r}{R} \tag{43}$$

After $\Delta t=t_2-t_1$, i.e., at $t=t_2$:

$$\sin(\beta+\Delta\beta) = \frac{r}{R-|t|\Delta t} \tag{44}$$

For infinitesimally small Δt and by using $\lim_{x\to 0}\frac{1}{1-x} = 1+x$ the last equation becomes

$$\sin(\beta+\Delta\beta) = \frac{r}{R}\,(1+\frac{|t|}{R}\Delta t) \tag{45}$$

By subtracting Equation (43) from Equation (45), then dividing by Δt, and letting

$\Delta t \rightarrow 0$ we obtain:

$$\frac{d\sin\beta}{dt} = \frac{r\,|t|}{R^2} \tag{46}$$

Using the following relations: $\frac{r}{R} = \sin\beta$, $\frac{|t|}{R} = -\frac{\frac{dR}{dt}}{R}$ and $\frac{d\sin\beta}{dt} = \cos\beta\frac{d\beta}{dt}$ in Equation (46) we obtain:

$$\frac{\frac{dR}{dt}}{R} = -\frac{\frac{d\beta}{dt}}{\tan\beta} \tag{47}$$

and the expression for looming becomes:

$$L = -\frac{\dot{R}}{R} = \frac{\dot{\beta}}{\tan\beta} = \frac{d}{dt}\ln(\sin\beta) \tag{48}$$

where dot denotes derivative with respect to time.

This means that the looming can be measured using β and $\dot{\beta}$ relative to the center of the ball. $\frac{\dot{\beta}}{\tan\beta}$ contains the information on the relative expansion of the projected ball during fixation at the center of the ball. This also means (as detailed later) that a retina which is constructed in a logarithmic fashion, i.e., the pixel length $\Delta\beta$ is proportional to $\tan\beta$, will measure looming in a linear fashion independently of the size of the object (as long as the retina fixates at the center of the object).

The above derivation of looming is based on a 2-D angle measurement of a projection of a ball. However, the extension to the projected area of a ball is simple. Relative change of area is also proportional to looming. We shall show next that

$$\frac{\frac{dA}{dt}}{2A} \approx -\frac{\frac{dR}{dt}}{R} = L \tag{49}$$

where A is the projected area of the ball on an spherical image.

The area occupied by a projected ball on a spherical image is $A = 4\pi\sin^2(\frac{\beta}{2})r_1^2$ where β is as defined in Figure 12, and r_1 is the radius of the sphere of the retina. r_1

is constant. By computing $\frac{\frac{dA}{dt}}{2A}$ which is the relative change of the projected area (divided by 2), we obtain: $\frac{\frac{dA}{dt}}{2A} = \frac{\frac{d\beta}{dt}}{2\tan\frac{\beta}{2}}$. For small β, $\frac{\frac{d\beta}{dt}}{2\tan\frac{\beta}{2}} \approx \frac{\frac{d\beta}{dt}}{\tan\beta}$ which has been shown (Equation (48)) to be the expression for looming. Thus the relative change of the projected area can be used to approximate looming values using Equation (49).

6.2.2 THE 2-D PLATE SCENARIO

We describe how to measure looming of a flying 2-D circular plate undergoing an arbitrary six-degree-of-freedom motion. Figure 13 shows a sequence of images of a two-dimensional plate. Note that the change in the projected area cannot be used to measure looming. However, at each instant of time there is a plane which is perpendicular to the line that connects the pinhole point of the camera with the center of the plate (Figure 14). The intersection between the plane and the plate is a line (except for one singular case where the intersection is a circle) whose projection is visible to the camera. Since this line (which is also a diameter of the plate) is perpendicular to the line that connects the camera with the plate, it will almost have the longest projection (compared to all other diameters' projections). The projection of this line can be measured. The relative change in length of this line can be used as a measure for looming. Similar computations as described previously for a section of a ball hold here.

6.2.3 OTHER OBJECTS

The problem with this approach is that in many cases it is difficult or even impossible to measure looming using the described method since a real environment is not filled with 3-D balls or circular plates. In several cases, however, it is possible to get good approximations of looming especially when points lie near the plane which is perpendicular to the optical axis.

6.3 QUALITATIVE MEASUREMENT OF LOOMING

Looming can be measured in qualitative terms. This means that instead of measuring the exact value of looming, we may be interested if the looming value is

above or below a threshold level. Alternatively, we may be interested to check if the looming value is within a specific margin. Figure 15 shows what we mean by qualitative looming (see also [33]). In Figure 15(a), 3-D space is divided into two regions. One consists of all points inside an equal looming sphere, and the second region corresponds to all points on or outside the equal looming sphere. Points inside this sphere are considered to be in a "Danger Zone" since they produce "above threshold" values of looming, while points on or outside the ball are in a "Safety Zone" since they produce "below threshold" values of looming.

A more quantitative division of 3-D space is shown in Figure 15(b). Here the space is divided into several regions, according to the range of looming values of points in each region. We called these regions "Danger Zone", "High Risk Zone", and "Safety Zone".

6.4 ON THE MEASUREMENT OF LOOMING WITH A LOGARITHMIC SENSOR

In this section we show that it is "natural" to use a logarithmic sensor to measure looming. (The results are highly related to neurobiologists' observations about retinal mapping [61]).

What is a "logarithmic sensor"? If the camera has a multiresolution imaging chip, where the pixel's radial size is approximately proportional to its distance from the optical axis, then it is called "logarithmic". Figure 16 shows a log-polar retina. This is different from a linear-polar retina (Figure 17) where the pixel radial length is constant. (We wrote "approximately proportional" since it is not possible to measure uniquely the distance from a finite length pixel to the optical axis.) Similarly, for a spherical image, if the pixel's angular size is proportional to the angle (or a function of the angle e.g., sin(.)) between the optical axis and the line that connects the center of the sphere with the pixel, it is also called a multiresolution logarithmic retina.

The relevant mathematical expression for the spherical retina (for the continuous case) where the resolution is proportional to the angle from the optical axis is:

$$d\ln\beta = \frac{d\beta}{\beta} \tag{50}$$

where β is the angle between the optical axis and the line that connects the center of the sphere with a particular point on the image sphere. In practice since pixel size $\Delta\beta$ is finite this expression becomes:

$$\Delta \ln\beta \approx \frac{\Delta\beta}{\beta} \tag{51}$$

Another example for a logarithmic retina is where $d\beta$ is proportional to $\tan\beta$. In this case the relevant mathematical expression is:

$$d \ln(\sin\beta) = \frac{d\beta}{\tan\beta} \tag{52}$$

and when the pixel size $\Delta\beta$ is finite, Equation (52) becomes:

$$\Delta \ln(\sin\beta) \approx \frac{\Delta\beta}{\tan\beta} \tag{53}$$

We have shown in section 6.2.1 that for the 3-D ball scenario (given that the optical axis of the camera passes through the center of the ball) the looming can be obtained from the ratio $\frac{\dot{\beta}}{\tan\beta}$ (Equation (48)) independently of the camera rotation, where β is the angle between the center of the ball and a point on the circular edge of the ball. This means that by using the $\ln(\sin\beta)$ retina (Figure 16, as opposed to the linear retina of Figure 17) it is possible to measure looming from the expansion of the projection of a ball in a simple way even if rotation is involved (as long as the optical axis passes through the center of the ball, i.e., the camera "fixates" at the center of the ball). In other words, by constructing a retina as mathematically described in Equation (52) for the continuous case, or as in Equation (53) for the discrete case, then motion of a point in the image from pixel ring i to pixel ring $i+1$ corresponds to the same looming value as motion of a point in the image from pixel j to pixel $j+1$, and the same as n times *less* the looming of a change from pixel k to $k+n$, for any positive integers i, j, k, n. In other words, the $\ln(sin)$ *retina measures looming during fixation in a linear way!*

Note that the expression for looming of a 3-D ball (48) is equivalent to the expression obtained for optical flow (42) for points on 3-D equal looming spheres when

the camera undergoes translation only along the optical axis. (θ in Equation (42) is measured from the X axis which is perpendicular to the optical axis, and β in Equation (48) may be measured relative to the optical axis. Expressions (42) and (48) become identical when substituting $\beta = 90°-\theta$ and $\dot{\beta}=-\dot{\theta}$.) Also, the relative change in the object's size and the corresponding change in the pixels' number on the ln(*sin*) retina are proportional as can be seen from Equations (48) and (49)). So far the existing literature has shown how a logarithmic retina can be used for simplifying (i.e., linearizing) the measurement of the optical flow and time-to-contact from optical flow for a camera undergoing translation only [65,68,69]. Here we showed that the same retina can be used to simplify the measurement of looming of a point in translational motion and for fixation on the center of a ball. Recently ([78]), several other advantages of the logarithmic retina have been shown.

The derivation here is an extension of existing results and may be another advantage for using fixation. Using this particular logarithmic mapping, the measurement of looming becomes independent of the object's size, and looming becomes linearly mapped, i.e., different size balls located on the same equal looming surface result in the same *relative* change in pixels of the logarithmic retina. Figure 18 shows two different size balls which lie on the same equal looming sphere. Even though their projections on the retina are different, their *relative* rate of expansion is the same. Using a fixating logarithmic retina (Figures 19 and 20) (as described previously) their expansion in terms of number of pixels is the same.

7. LOOMING AS A FUNCTION OF TIME

So far we have dealt with instantaneous looming. We now examine looming as a function of time. As shown in Section 4.2 equal looming surfaces are spheres. The physical size of each sphere is proportional to the instantaneous translational velocity, the centers of the spheres lie on a line along the instantaneous velocity vector, and the observer lies on the sphere surface. Also, the spheres are independent of the instantaneous rotational parameters.

Figure 21 shows what happens to these spheres as a function of time. Note that their centers always lie on a line along the instantaneous translational vector and *not* on

the optical axis. Figure 22 examines these spheres for a particular point "A" in space. Assume a camera moving in constant speed along a straight line. As the camera approaches the point "A" (for $t<t_4$), the equal looming spheres containing this point shrink i.e., produce higher values of looming. At a $\theta = 45°$ angle as in Figure 8 (i.e., at time instant $t=t_4$) the sphere containing point "A" gets its minimum size which means that the looming value is maximum. Then, when the camera translational vector is perpendicular to the line that connects the point "A" with the camera pinhole point (time instant $t=t_5$), the sphere becomes a plane, i.e., the point "A" produces zero value of looming. At $\theta = -45°$ (at $t=t_6$) the looming value is minimum, and as the camera continues to move ($t>t_6$)the looming value approaches zero.

Figure 23 shows the values of looming of several points as a function of time. Figure 23a shows a camera that undergoes both translation and rotation. We examine the looming values for four different points: 1,2,3, and 4. The looming values are plotted in Figure 23b. At time instant $t = t_1$ the looming values for all four points are positive, and the looming value of point 2 gets its maximum value. At $t = t_2$, the looming value of point 3 gets its maximum value since the angle formed by the instantaneous translation vector and the line that connects the pinhole point of the camera and point 3 is 45°. At time close to $t = t_3$ the looming produced by point 4 gets a very high positive value (which changes to a very high negative value at the time instant immediately after $t = t_3$). At $t = t_3$ points 1, 2, and 3 are all on the same equal looming sphere and point 1 produces a maximum looming value. Note that the looming values of points 1, 2, and 3 are the same at $t = t_3$ as shown in Figure 23b since they lie on an equal looming sphere.

Shortly after $t = t_3$ point 3 produces zero looming value. At $t = t_4$ point 3 produces minimum value of looming and point 2 produces zero looming value. At $t = t_5$ point 2 produces minimum value of looming and point 1 produces zero value of looming. At $t = t_6$ points 1, 2, 3, and 4 produce the same value of looming and point 1 produces minimum value of looming. Note that looming values are independent of the rotation of the camera. (The camera in Figure 23a was chosen to fixate on point 2.)

The case where a point moves relative to a camera with *constant* value of looming can be described in two basic ways. In Figure 24a a point is moving in the XY plane. At each instant of time, e.g., t_1, t_2, t_3, and t_4, the point is on a *physically* different sphere, but since the looming values produced by the point are the same for all time instants the point appears on the *same* equal looming sphere. In Figure 24b the coordinates are scaled by a factor of $\frac{1}{|t|}$. Note that the relative translational velocity of the point changes in order to keep a constant looming value at all time instants.

8. MORE ON LOOMING

8.1 LOOMING AND THE TIME-TO-CONTACT

As mentioned earlier, looming is different from the time-to-contact. It is related more to the "two dimensional" time-to-contact concept [21]. According to Lee [9] any point that lies on a plane which is perpendicular to the instantaneos translational motion direction of the camera, will produce the same value of "time-to-contact" T_c (assuming that the optical axis coincides with the direction of motion). This means that the time-to-contact deals with *depth*. The derivation in [9] is valid only for rectilinear motion (with no rotation) of the camera. Only recently [25,26] has it been extended to a more general motion of the camera.

One problem with the time-to-contact approach is that points which lie on a single perpendicular plane but are located far away from, or close to, the camera produce the same value of T_c even though they are not equally relevant to making vision-based behavioral decisions. The looming value of a point is related to *range* rather than depth.

Figures 25 and 26 illustrate the main difference between the time-to-contact value of a point, and its looming value. All points that lie on a vertical plane will have the same time-to-contact value. However, points on a sphere in front of the camera produce the same looming value. Points 1,2, and 3 in Figure 26 have the same looming value but different "time to contact" values. Points 2 and 4 have the same time-to-contact value but different looming values. Points 4 and 5 have the same time-to-contact values and the same looming values.

8.2 LOOMING, ZOOMING, AND VIEWING

Note that since looming deals with the *relative* change of range the looming values obtained from zooming or non-zooming cameras are the same, although the images are different in these two cases. Also the looming perceived by a camera (#1 in Figure 27) is the same as that perceived by other cameras (e.g., #2, #3, and #4 in Figure 27) that observe a monitor on which images from camera #1 are displayed (regardless of the distance and location relative to the monitor) This is due to the fact that the *relative* change in the area is the same from all points of view.

9. LOOMING AND THE EQUAL FLOW CIRCLES

9.1 THE EQUAL FLOW CIRCLES

This section deals with points in 3-D space that produce the same value of optical flow. For a particular motion of the camera point with a specific value of optical flow lie on a circle. We briefly review this result.

It has been shown in previous work [63] that if the camera motion vectors $\mathbf{t}$ and $\boldsymbol{\omega}$ are:

$$\mathbf{t} = (U,V,0)^T$$

and

$$\boldsymbol{\omega} = (0,0,C)^T.$$

Then the optical flow expression in spherical coordinates (Figure 11) is:

$$\begin{bmatrix}\dot{\theta}\\ \dot{\phi}\end{bmatrix} = \begin{bmatrix} \frac{-Y}{X^2+Y^2} & \frac{X}{X^2+Y^2} & 0 \\ \frac{-XZ}{\sqrt{X^2+Y^2}(X^2+Y^2+Z^2)} & \frac{-YZ}{\sqrt{X^2+Y^2}(X^2+Y^2+Z^2)} & \frac{\sqrt{X^2+Y^2}}{X^2+Y^2+Z^2} \end{bmatrix} \begin{bmatrix} -U+CY \\ -V-CX \\ 0 \end{bmatrix} \tag{54}$$

From Equation set (54), the points in space that result from constant $\dot{\theta}$ (regardless of the value of $\dot{\phi}$) form a cylinder of infinite height whose equation is $\left[X+\frac{V}{2(C+\dot{\theta})}\right]^2+\left[Y-\frac{U}{2(C+\dot{\theta})}\right]^2=\left[\frac{V}{2(C+\dot{\theta})}\right]^2+\left[\frac{U}{2(C+\dot{\theta})}\right]^2$, as displayed in Figure 28a, and the points in space that result from $\dot{\phi}=0$ (regardless of the value of $\dot{\theta}$) are those that

lie on (a) a plane whose equation is $Y = -\frac{U}{V}X$, or (b) a plane whose equation is $Z=0$ (i.e., the XY plane) as pictorially described in Figure 28b. The intersection of the cylinder with the planes is the desired solution (Figure 28c), i.e., the points in space that result in $\dot{\theta} = constant$ *and* $\dot{\phi} = 0$ optical flow values. Analytically, the following are the solutions (disallowing the case of $X = 0$ and $Y = 0$ which corresponds to an anomalous situation):

$$Z = 0 \text{ and } \left[X + \frac{V}{2(C+\dot{\theta})}\right]^2 + \left[Y - \frac{U}{2(C+\dot{\theta})}\right]^2 = \left[\frac{V}{2(C+\dot{\theta})}\right]^2 + \left[\frac{U}{2(C+\dot{\theta})}\right]^2. \quad (55)$$

$$X = -\frac{V}{C+\dot{\theta}} \text{ and } Y = \frac{U}{C+\dot{\theta}}. \quad (56)$$

These solutions are drawn in Figure 29. Solution (55) is an equation of a circle that lies in the XY plane. The radius of the circle is $\left[\left[\frac{V}{2(C+\dot{\theta})}\right]^2 + \left[\frac{U}{2(C+\dot{\theta})}\right]^2\right]^{\frac{1}{2}}$ and its center is at $\left[-\frac{V}{2(C+\dot{\theta})}, \frac{U}{2(C+\dot{\theta})}\right]$. The circle is tangent to the camera translation vector at the origin.

Solution (56) is a straight line perpendicular to the XY plane and intersects this plane at the point $\left[-\frac{V}{C+\dot{\theta}}, \frac{U}{C+\dot{\theta}}, 0\right]$. This intersection point also lies on the circle defined in solution (55).

The meaning of these solutions is the following: all points in 3-D space that lie on the circle or the line described by solutions (55) and (56) and which are visible (i.e., unoccluded and in the field of view of the camera) produce the same instantaneous optical flow $\dot{\theta}$ and zero instantaneous optical flow $\dot{\phi}$. We call the circle on which equal flow points lie the Equal Flow Circle (EFC). Two sets of EFCs are illustrated in Figure 30. Figure 30a shows EFCs for the case where the camera undergoes instantaneous translation only. The label of each circle represents the optical flow $\dot{\theta}$ in the image that corresponds to points on this circle. Here, there is a straight line (a circle with an infinite radius) that corresponds to zero flow in the image. Figure 30b shows EFCs for

the case where the camera undergoes both instantaneous translation and rotation. Here, there is a circle with finite radius that produces zero flow ($\dot{\theta} = 0$ in the image domain). We called this circle the Zero Flow Circle (ZFC).

9.2 THE EFCs and ZFCs AS A FUNCTION OF TIME

As the camera moves through 3-D space, the EFCs move with it. Figure 31 is an example of a camera path with some EFCs. At each instant of time, the radii of the EFCs are a function of the instantaneous motion parameters $\mathbf{t}$ and $\boldsymbol{\omega}$. The locations of the EFCs are such that they always contain the origin of the camera coordinate system (the same as the camera pinhole point), are tangent to the instantaneous translation vector $\mathbf{t}$, and are perpendicular to the instantaneous rotation vector $\boldsymbol{\omega}$. Each ZFC lies to the left or right of the translation vector depending on whether the instantanous rotation is positive or negative, respectively.

9.3 LOOMING - EFCs MAPPING

As described above, there are equal looming circles in the $X-Y$ plane which are sections of the equal looming spheres. These circles are orthogonal to the EFCs. This observation suggests a 2-D mapping that consists of two orthogonal families of circles as shown in Figure 32.

A point in 2-D space can be mapped in camera coordinates by specifying two numbers. One is the optical flow of the point, and the other is the looming value of the point.

Figure 33 and 34 show how a segment in the new domain (segment 11,12,22,21 in Figure 33) can be mapped to another orthogonal grid (segment 11,12,22,21 in Figure 34). In Figure 34 the horizontal axis corresponds to the optical flow, and the vertical axis corresponds to the value of looming. Note that the vertical line for which $\dot{\theta} = 0$ corresponds to the Zero Flow Circle.

Regions for obstacle avoidance and other behavior related tasks can be defined using the looming-EFCs domain. In Figure 35 a qualitative partition of the 2-D plane is shown. A region called "danger zone" can be used to qualitatively detect obstacles. A more quantitative partition is shown in Figure 36 where space is divided into "danger

zone", "high risk zone", and "safety zone".

Clearly, the mapping can be extended to 3-D using, for example, equal flow cylinders and equal looming spheres.

10. DISCUSSION

In this paper we presented a new and mainly quantitative approach for analyzing looming. Using the concept of the equal looming spheres we showed how space can be perceived during motion. Several methods for measuring looming of simple 3-D objects were presented. It has been shown that using a logarithmic retina the looming computations become simpler. We also showed how beneficial fixation can be to simplify the measurement of looming. This analysis complements the qualitative understanding of looming and fixation.

The equal looming spheres can be thought of as properties of space rather than properties of points in the image domain.

The coordinate system that we chose is a convenient one. However, the angular velocities of points in space are independent of their representation in the image. Other image coordinate systems may be chosen, in particular, (for practical purposes) the image domain may be planar. (Obviously, an appropriate conversion from the spherical coordinate system should be used).

The approach for measuring looming using relative rate of expansion is quite limited. It works mainly for sphere and sphere-like objects, or sections of these objects.

Computing looming from image flow by the method of Horn and Schunck [36] has the advantage of simplicity. Unfortunately, it tends to produce noisy and inaccurate results. There are four sources of noise and errors: First, the sampled data system approximations to $\frac{dI}{dt}$, $\frac{dI}{d\theta}$, and $\frac{dI}{d\phi}$, are subject to digitization noise and Nyquist sampling frequency limitations. Second, except in the vicinity of brightness edges, $\frac{dI}{dt}$ is small, and may be zero. Division by small numbers magnifies errors. Third, the sensitivity of photodetectors in any array is not uniform. The difference in signal from two adjacent photo-detectors is thus not necessarily an accurate measure of the

difference in illumination. Thus range from image flow by the Horn and Schunck method tends to be inaccurate for smooth images, and unreliable for sharply focused natural images.

Computing looming from image flow by the method of cross-correlation [53] can be much more accurate and noise free. It is, however, more complicated and computationally intensive.

11. FUTURE WORK

Future work will focus on measures of performance of the above algorithms in an environment filled with a variety of natural objects. Tests will be made that simulate tasks such as obstacle avoidance. We plan to exploit the equal looming surfaces concept in a vision based navigation algorithm for a real-time robot system.

12. ACKNOWLEDGEMENTS

The author would like to thank Jim Albus, Ernie Kent, Marty Herman and Don Orser for fruitful discussions regarding looming. Jim Albus and Ernie Kent have continuously encouraged discovering and establishing vision related computational theories. Hakan Yakali produced the high quality figures that appear in this paper.

REFERENCES

[1] Alderson, G.J.K., Sully, D.L. and Sully, H.G., "An operational analysis of a one-handed catching task using high speed photography", *Journal of Motor Behavior,* 6, pp. 217-226, 1974.

[2] Beek, P.J., "Perception-action coupling in the young infant: An appraisal of von Hofsten's research programme", In M.G. Wade and H.T.A. Whiting (Eds.), *Motor development in children: Aspects of coordination and control,* Dordrecht, The Netherlands: Martinus-Nijhoff, pp. 187-196, 1986.

[3] Bootsma, R.J., *The timing of rapid interceptive actions,* Amsterdam: Free University Press, 1988.

[4] Bootsma, R.J., "Accuracy of perceptual processes subserving different perception-action systems", *Quarterly Journal of Experimental Psychology,* 41A, pp. 489-500, 1989.

[5] Bootsma, R.J. and van Wieringen, P.W.C., "Timing an attacking forehand drive in table tennis", *Journal of Experimental Psychology: Human Perception and Performance,* 16, pp. 21-29, 1990.

[6] Bower, T.G.R., Broughton, J.M. and Moore, M.K., "The coordination of visual and tactual input in infants", *Perception & Psychophysics,* 8, pp. 51-53, 1970.

[7] Jeannerod, M., "The timing of natural prehension movements", *Journal of Motor Behavior,* 16, pp. 235-254, 1984.

[8] Laurent, M., Dinh Phung, R. and Ripoll, H., "What visual information is used by riders in jumping", *Human Movement Science,* 8, pp. 481-501, 1989.

[9] Lee, D.N., "A theory of visual control of braking based on information about time-to-collision", *Perception,* 5, pp. 437-459, 1976.

[10] Lee, D.N., "Visuo-motor coordination in space-time", In G.E. Stelmach and J. Requin (Eds.), *Tutorials in motor behavior,* Amsterdam: North-Holland, pp. 281-293, 1980.

[11] Lee, D.N., Lishman, J.R. and Thomson, J.A., "Regulation of gait in long jumping", *Journal of Experimental Psychology: Human Perception and Performance,* 8, pp. 448-459, 1982.

[12] Lee, D.N. and Reddish, D.E., "Plummeting gannets: A paradigm of ecological optics", *Nature,* 293, pp. 293-294, 1981.

[13] Lee, D. N. and Young, D.S., "Visual timing of interceptive actions", In D.J. Ingle, M. Jeannerod, and D.N. Lee (Eds.), *Brain mechanisms and spatial vision,* Dordrecht, The Netherlands: Martinus Nijhoff, pp. 1-30, 1985.

[14] Lee, D.N., Young, D.S., Reddish, D.E., Lough, S. and Clayton, T.M.H., "Visual timing in hitting an accelerating ball", *Quarterly Journal of Experimental Psychology,* 35A, pp. 333-346, 1983.

[15] Marteniuk, R.G., Mackenzie, C.L. and Leavitt, J.L., "The inadequacies of a straight physical account", In H.T.A. Whiting, O.G. Meijerand P.C.W. van Wieringen (Eds.), *The natural/physical approach to movement control,* Amsterdam: Free University Press, pp. 95-115, 1990.

[16] McLeod, P., McLaughlin, C., and Nimmo-Smith, I., "Information encapsulation and automaticity: Evidence from the visual control of finely-timed actions", In M. Posner & O. Malin (Eds.), *Attention and performance XI,* Hillsdale, NJ: Erlbaum, pp. 391-406, 1986.

[17] McLeon, R.W. and Ross, H.E., "Optic-flow and cognitive factors in time-to-collision estimates", *Perception,* 12, pp. 417-423, 1983.

[18] Regan, D.M., "The eye in ball games: Hitting and catching", In *Proceedings of the Conference on Vision and Sport,* Harlem, The Netherlands: De Vriesenborch, pp. 1-33, 1986.

[19] Schiff, W., "Perception of impending collision: A study of visually directed avoidance behavior", *Psychological Monographs,* 79 (Whole No. 604), 1965.

[20] Schiff, W., Caviness, J.A. and Gibson, J.J., "Persistent fear responses in rhesus monkeys to the optical stimulus of 'looming'", *Science,* 136, pp. 982-983, 1962.

[21] Schiff, W., and Detweiler, M.L., "Information used in judging impending collision", *Perception,* 8, pp. 647-658, 1979.

[22] Savelsbergh, G.J.P., Whiting, H.T.A. and Bootsma, R.J., "Grasping Tau", *Journal of experimental Psychology,* Vol. 17, no. 2, pp. 315-322, 1991.

[23] Todd, J.T., "Visual information about moving objects", *Journal of Experimental Psychology: Human Perception and Performance,* 7, pp. 795-810, 1981.

[24] Wallace, S.A., Weeks, D.L., and Kelso, J.A.S., "Temporal constraints in reaching and grasping behavior", *Human Movement Science,* 9, pp. 69-93, 1990.

[25] Balasubramanyan, P. and Snyder, M.A., "The P-Field: A Computational Model for Binocular Motion Processing", *In Proc. IEEE Computer Vision and Pattern Recognition (CVPR) Conference,* Maui, Hawaii, June, 1991.

[26] Tistarelli, M., Grosso, E. and Sandini, G., "Dynamic Stereo in Visual Navigation", *In Proc. IEEE Computer Vision and Pattern Recognition (CVPR) Conference,* Maui, Hawaii, June, 1991.

[27] Raviv, D., "Extraction of the 'Time To Contact' from Real Visual Data", *In Proc. SPIE's Symposium on Advances in intelligent Robotics Systems* (1194), Pennsylvania, November 1989.

[28] Sandini, G. and Tistarelli, M., "Robust obstacle detection using optical flow", *In Proc. of IEEE Intl. Workshop on Robust Computer Vision,* pp. 396-411, Seattle, WA, October, 1990.

[29] Crowley, J.L., "Dynamic modeling of free-space for a mobile robot", *In Proc. of IEEE/RJS Intl. Workshop on Intelligent Robots and Systems,* Tsukuba, Japan, September, 1989.

[30] Nelson, R.C., and Aloimonos, J., "Using flow field divergence for obstacle avoidance in visual navigation", *IEEE Trans. on PAMI,* PAMI-11, No. 10, October 1989.

[31] Ballard, D.H., Nelson, R.C. and Yamauchi, B., "Animate vision", *Optics News,* 15(5), pp. 17-25, 1989.

[32] Grosso, E., Sandini, G. and Tistarelli, M., "3d object reconstructiion using stereo and motion", *IEEE Trans. of Syst. Man and Cybernetics,* SMC-19, No. 6, November/December 1989.

[33] Aloimonos, J., "Purposive and qualitative active vision", In *Proc. Intl. Workshop on Active Control in Visual Perception,* Antibes, France, April, 1990.

[34] Sandini, G. and Tistarelli, M., "Active tracking strategy for monocular depth inference over multiple frames", *IEEE Trans. on PAMI,* PAMI-12, No. 1, January 1990.

[35] Aloimonos, J., Weiss, I. and Bandyopadhyay A., "Active vision", *Intl. Journal of Computer Vision,* 1(4) pp. 333-356, 1988.

[36] Horn, B.K.P. and Schunck B.G., "Determining optical flow", *Artificial Intelligence,* 17, pp. 185-204, 1981.

[37] Uras, S., Girkosi, F., Verri and Torre, V., "Computational approach to motion perception", *Biological Cybernetics,* 1988.

[38] Tistarelli, M. and Sandini, G., "Estimation of depth from motion using an anthropomorphic visual sensor", *Image and Vision Computing* 8, No. 4, pp. 271-278, 1990.

[39] Adiv G., "Determining Three-Dimensional Motion and Structure from Optical Flow Generated by Several Moving Objects," *IEEE PAMI,* Vol. 7, pp. 384-401, July 1985.

[40] Anandan, P., "A Computational Framework and Algorithm for the Measurement of Visual Motion," *IJCV,* Vol. 2, pp. 283-310, 1989.

[41] Balasubramanyam, P., "Computation of Motion-in-Depth Parameters Using Stereoscopic Motion Constraints" *Proc. IEEE Computer Society Workshop on Computer Vision,* Miami, pp. 349-351, 1987.

[42] Regan, D. and Beverley, K.I., "Binocular and Monocular Stimuli for Motion-In-Depth: Changing Disparity and Changing Size Inputs Feed the Same Motion-In-Depth Stage," *Vision Research,* 19, pp. 1331-1342, 1979.

[43] Longuet-Higgins H.C. and Pradzny, K., "The Interpretation of a Moving Retinal Image," *Proceedings of the Royal Society of London,* B(208):385-397, July 1980.

[44] Helmholtz, H. von, *Treatise on Physiological Optics* (translated by J.P.C. Southall), Dover Publications, New York, 1925.

[45] Gibson, J.J., *The Ecological Approach to Visual Perception,* Cornell University Press, Ithaca, N.Y., 1966.

[46] Rogers, B., and Graham, M., "Motion parallax as an independent cue for depth", *Perception* 8, pp. 125-134, 1979.

[47] Longuet-Higgins, H.C., and Prazdny, K., "The interpretation of a moving retinal image", *Proc. Royal Society of London,* B(208), pp. 385-397, 1980.

[48] Hildreth, E.C., *The Measurement of Visual Motion,* Ph.D. Thesis, MIT, August 1983.

[49] Waxman, A.M., and Wohn, K., "Image flow theory: a framework for 3-D inference from time-varying imagery", in *Advances in Computer Vision,* Vol. 1 (C. Brown, Editor), Lawrence Erlbaum Associates Publishers, 1988.

[50] Waxman, A.M., Kamgar-Parsi, B., and Subbarao, "Closed form solution to image flow equations", *International Journal of Computer Vision,* 1, pp. 239-258, 1987.

[51] Horn, B.K.P., *Robot Vision,* MIT Press, 1986.

[52] Meriam, J.L., *Dynamics,* John-Wiley & Sons, 1975.

[53] Albus, J.S., and Hong, T.H., "Motion, Depth, and Image Flow", *Proceedings IEEE International Conference on Robotics and Automation,* Cincinnati, OH, 1990.

[54] Raviv, D., "Parallel Algorithm for 3D Surface Reconstruction", *Proceedings, SPIE Symposium on Advances in Intelligent Robotics Systems,* Philadelphia, PA, Nov. 1989. Also NISTIR 1991.

[55] Rangachar, R., Hong, T., and Herman M., "Real-Time Differential Range Estimation Based on Time-Space Imagery Using PIPE", *Proceedings, SPIE Symposium on Image Processing II: Algorithms, Architectures, and Applications,* Vol. 1295, Orlando, FL, April 1990.

[56] Albus S.J., Internal note on fixation, NIST, 1989.

[57] Bajczy, R., "Active Perception", *Proceedings of the IEEE,* Vol. 76, No. 8, August 1988.

[58] Ballard, D., and Ozcandarli, A., "Eye Fixation and Early Vision: Kinetic Depth", *International Conference on Computer Vision,* pp. 524-531, 1988.

[59] Cutting, J.E., *Perception with an Eye for Motion,* MIT Press, 1986.

[60] Marr D., *Vision,* MIT Press, 1982.

[61] Schwartz, E., "Computational Anatomy and Functional Architecture of Striate Cortex: A Spatial Mapping to Perceptual Coding", *Vision Research,* Vol. 20, pp. 645-699, Pergamon Press, 1980.

[62] Raviv, D., and Herman, M., "Towards an Understanding of Camera Fixation", NIST Internal Report, NISTIR 89-4217; also in *Proceedings of IEEE Conference on Robotics and Automation,* Cincinnati, OH, 1990.

[63] Raviv, D., "A Quantitative Approach to Camera Fixation", In *Proc. Computer Vision and Pattern Recognition (CVPR) 91,* Hawaii, June 91.

[64] Raviv, D., and Albus J.S., "A Closed-Form Massively-Parallel Range-From-Image-Flow Algorithm", Accepted for publication *IEEE Transactions on Systems, Man and Cybernetics* March, 1992.

[65] Sandra Bartlett and Ramesh Jain, "Depth Determination Using complex Logarithmic Mapping", In *Proc. SPIE vol. 1382, Intelligent Robots and Computer Vision IX,* 1990.

[66] D. Ballard, "Animate Vision", *Artificial Intelligence* 48, pp. 57-86, 1991.

[67] Koendrick, J.J., "Scale Time", *Biological Cybernetics,* 58, pp. 159-162, 1988.

[68] G. Chaikin and Carl Weiman, "Log Spiral Grids In Computer Pattern Recognition", *Computer Graphics and Pattern Recognition,* Vol. 4, pp. 197-226, 1979.

[69] L. Massone, G. Sandini and v. Tagliasco, "Form-Invariant- Topological Mapping Strategy for 2-D Shape Recognition", *Computer Vision, Graphics and Image Processing,* Vol. 30, pp. 169-188, 1985.

[70] Weng, J., Huang, T.S., and Ahuja, N., "3-D Motion Estimation, Understanding, and Prediction from Noisy Image Sequences", *IEEE Transactions on Pattern Recognition and Machine Intelligence,* Vol. PAMI-9, No. 3, pp. 370-389, May 1987.

[71] Tsai, R.Y. and Huang, T.S., "Uniqness and Estimation of Three Dimentional Motion Parameters of Rigid Objects with Curved Surfaces", *IEEE Transactions on Pattern Recognition and Machine Intelligence,* Vol. PAMI-6, No. 1, pp. 13-26, Jan. 1984.

[72] Fang, J.Q., and Huang, T.S., "Some Experiments on Estimating the 3-D Motion Parameters of a Rigid Body from Two Consecutive Image Frames", *IEEE Transactions on Pattern Recognition and Machine Intelligence,* Vol. PAMI-6, pp. 545-554, 1984.

[73] Broida, T.J. and Chellappa, "Estimation of Object Motion Parameters from Noisy Images", *IEEE Transactions on Pattern Recognition and Machine Intelligence,* Vol. PAMI-8, No. 1, pp. 90-99, Jan. 1986.

[74] Bharwani, S., Riseman, E., and Hanson, A., "Refinement of Environment Depth Maps over Multiple Frames", In *Proceedings of the Workshop on Motion: Representations and Analysis,* Charleston, NC, pp. 73-80, May 1986.

[75] Horn, B.K.P. and Weldon, E.J., "Direct Methods for Recovering Motion", *International Journal of Computer Vision,* 2, pp. 51-76, 1988.

[76] Negahdaripour, S., and Horn, B.K.P., "Direct Passive Navigation", *IEEE Transactions on Pattern Recognition and Machine Intelligence,* Vol. PAMI-9, No. 1, pp. 641-650, December 1988.

[77] Nelson, R. and Aloimonos, J., "Using Flow Field Divergence for Obstacle Avoidance in Visual Navigation", DARPA Image Understanding (IU) Workshop, 1988.

[78] Tistarelli, M. and Sandini, G., "On the Advantages of Polar and Log-Polar Mapping for Direct Estimation of Time-To-Contact from Optical Flow", Technical report LIRA-TR3/90, University of Genoa, Italy, June 1990.

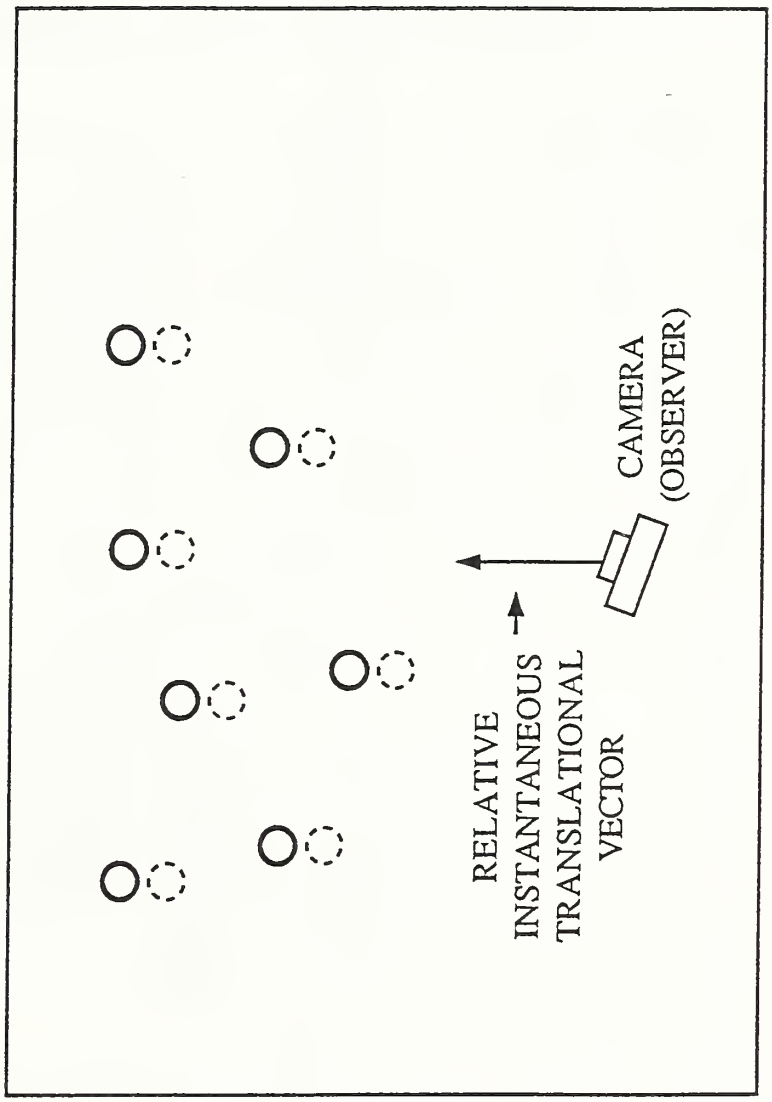

Figure 1: Relative Motion

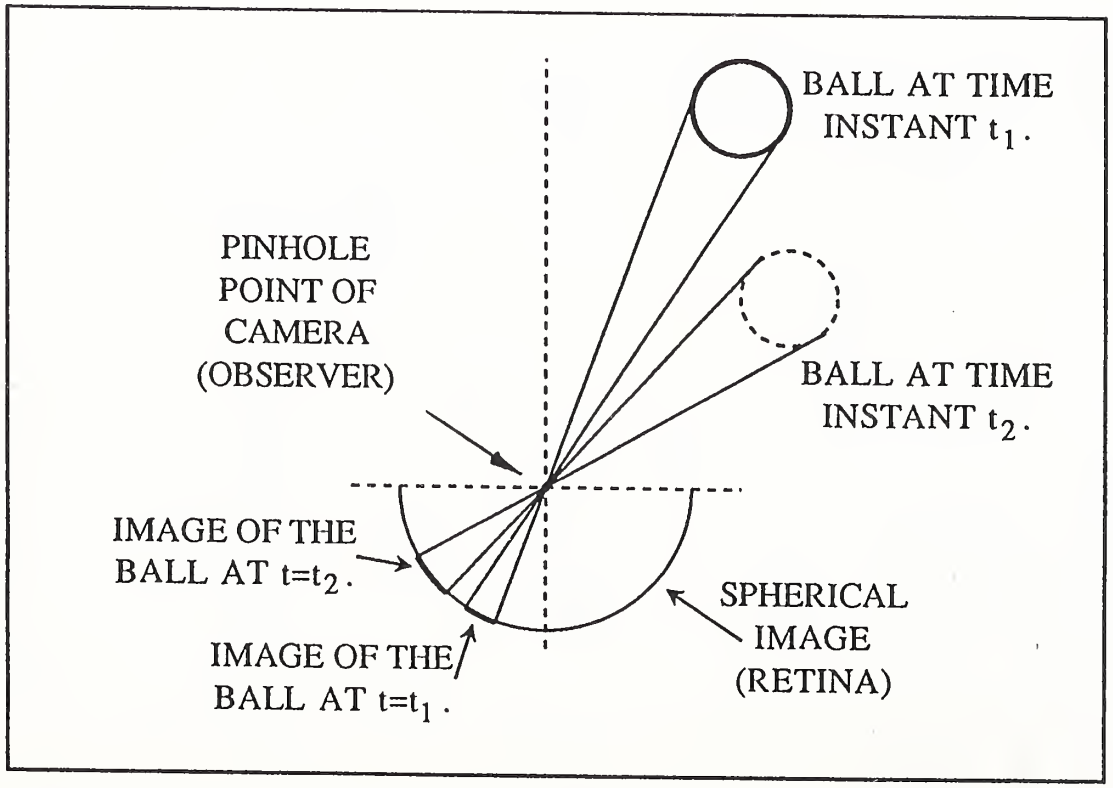

Figure 2(a): Projection of a Ball at Two Time Instants

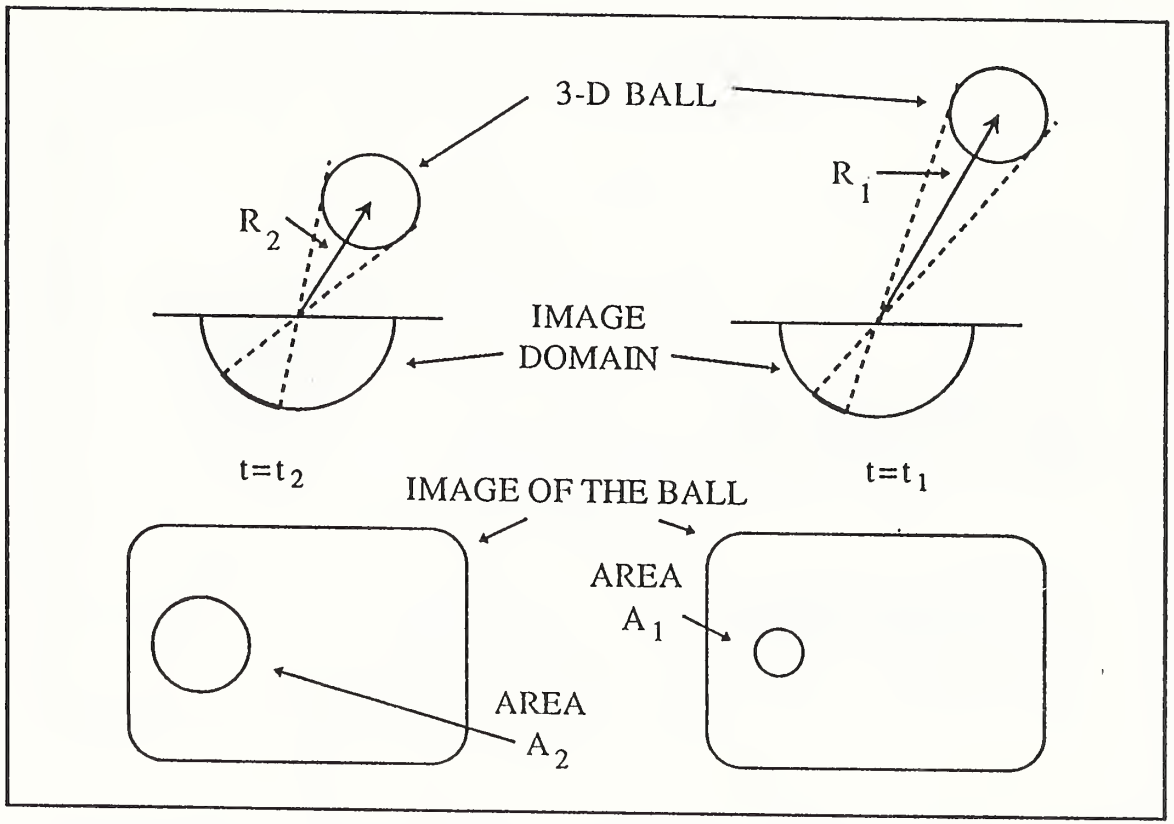

Figure 2(b): Projection of a Ball at Two Time Instants

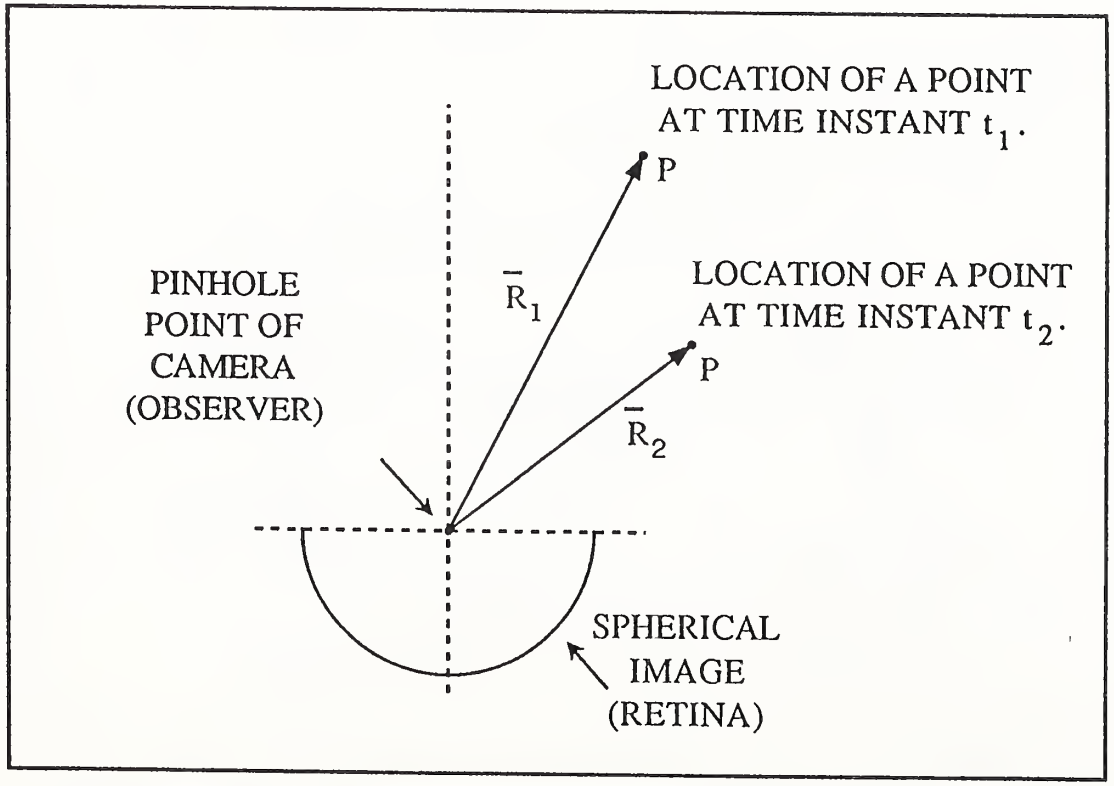

Figure 3: Motion Relative to a Point

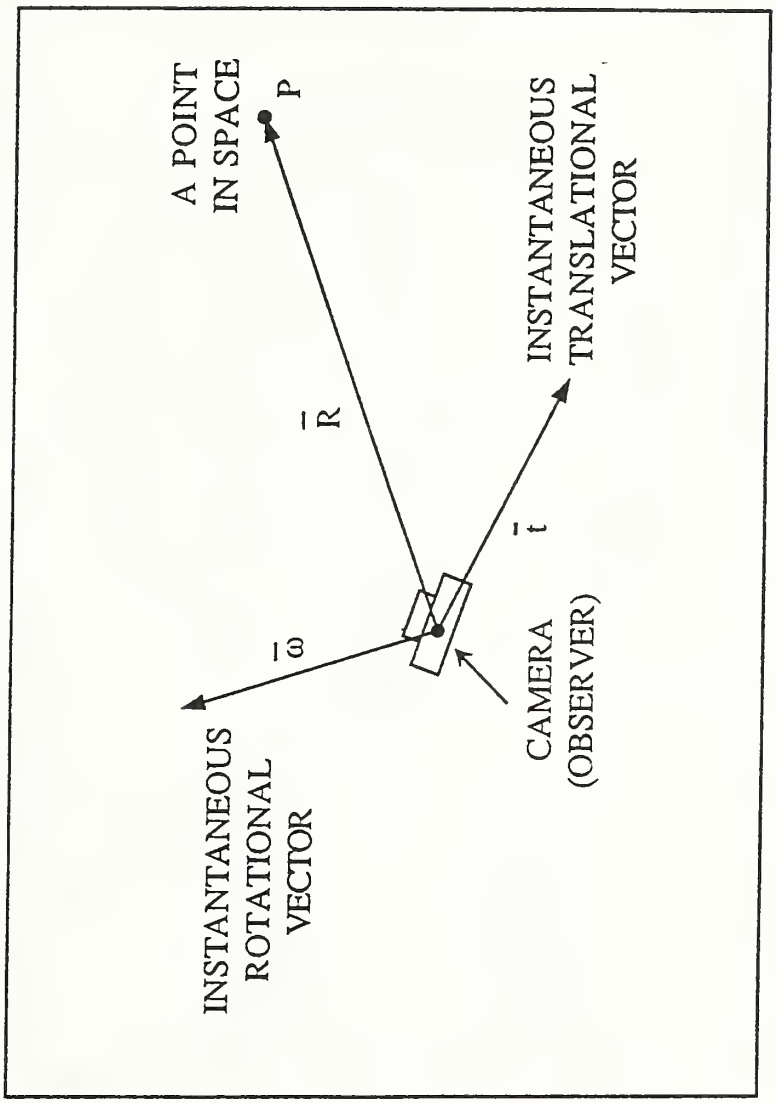

Figure 4: Motion Parameters

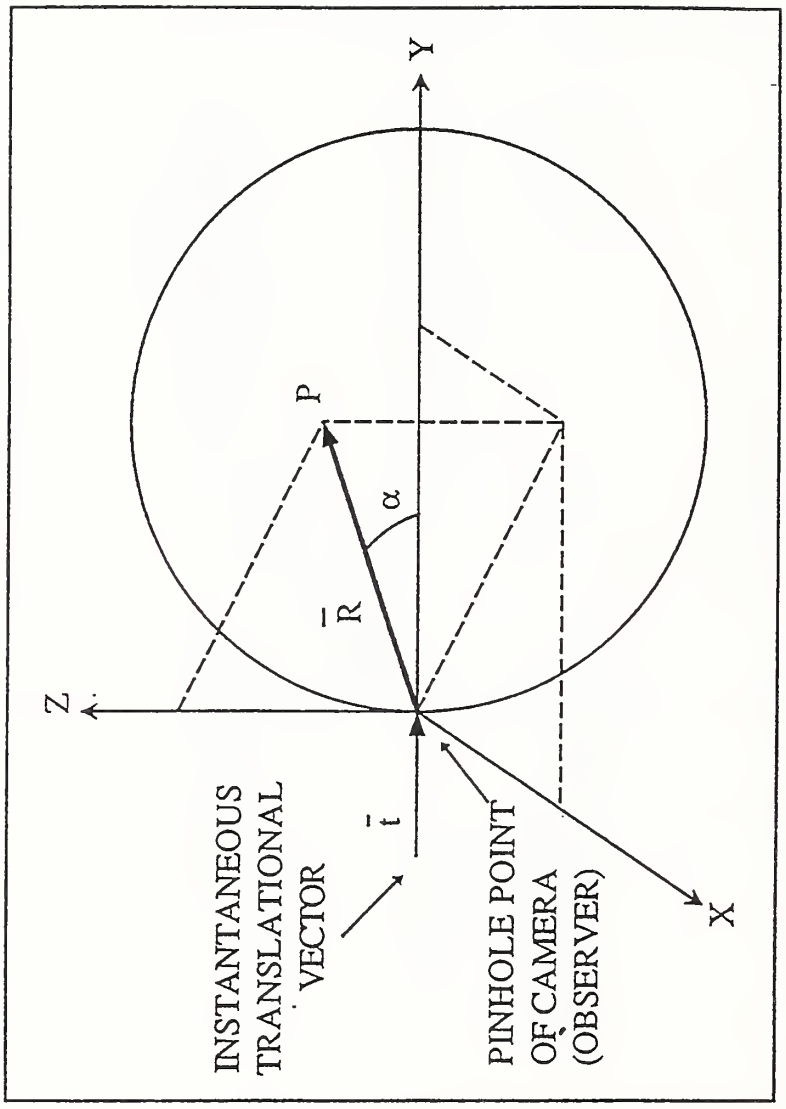

Figure 5: Equal Looming Sphere

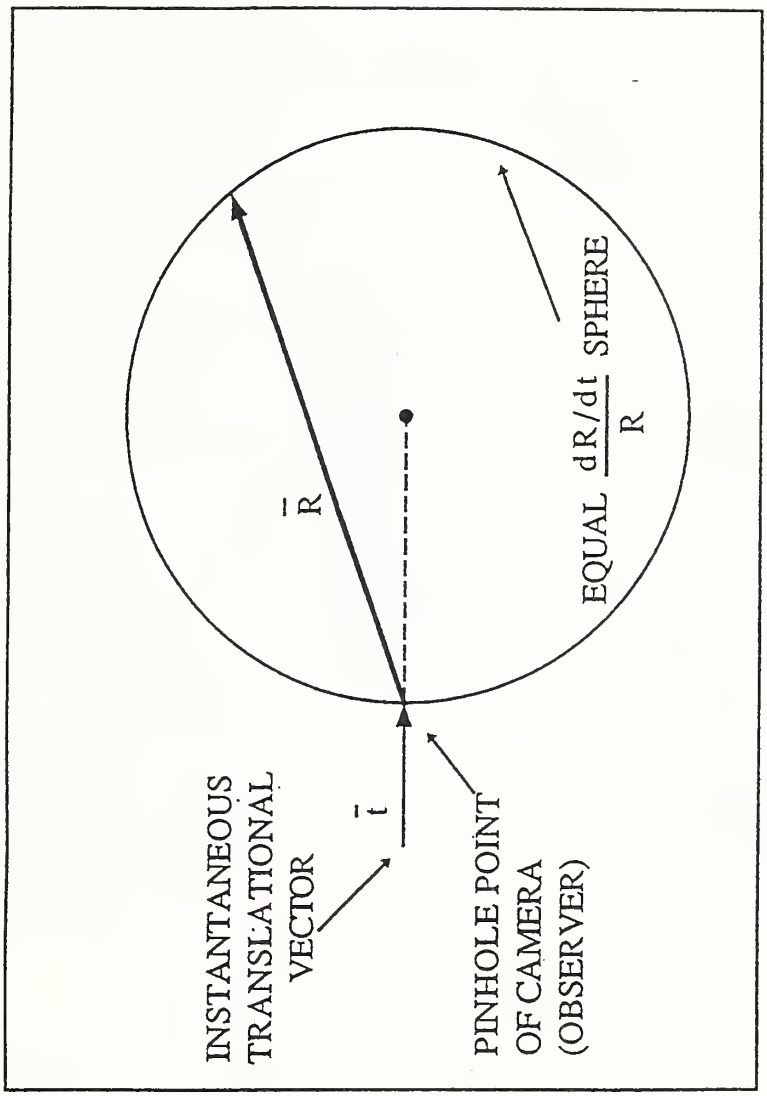

Figure 6: Equal Looming Sphere

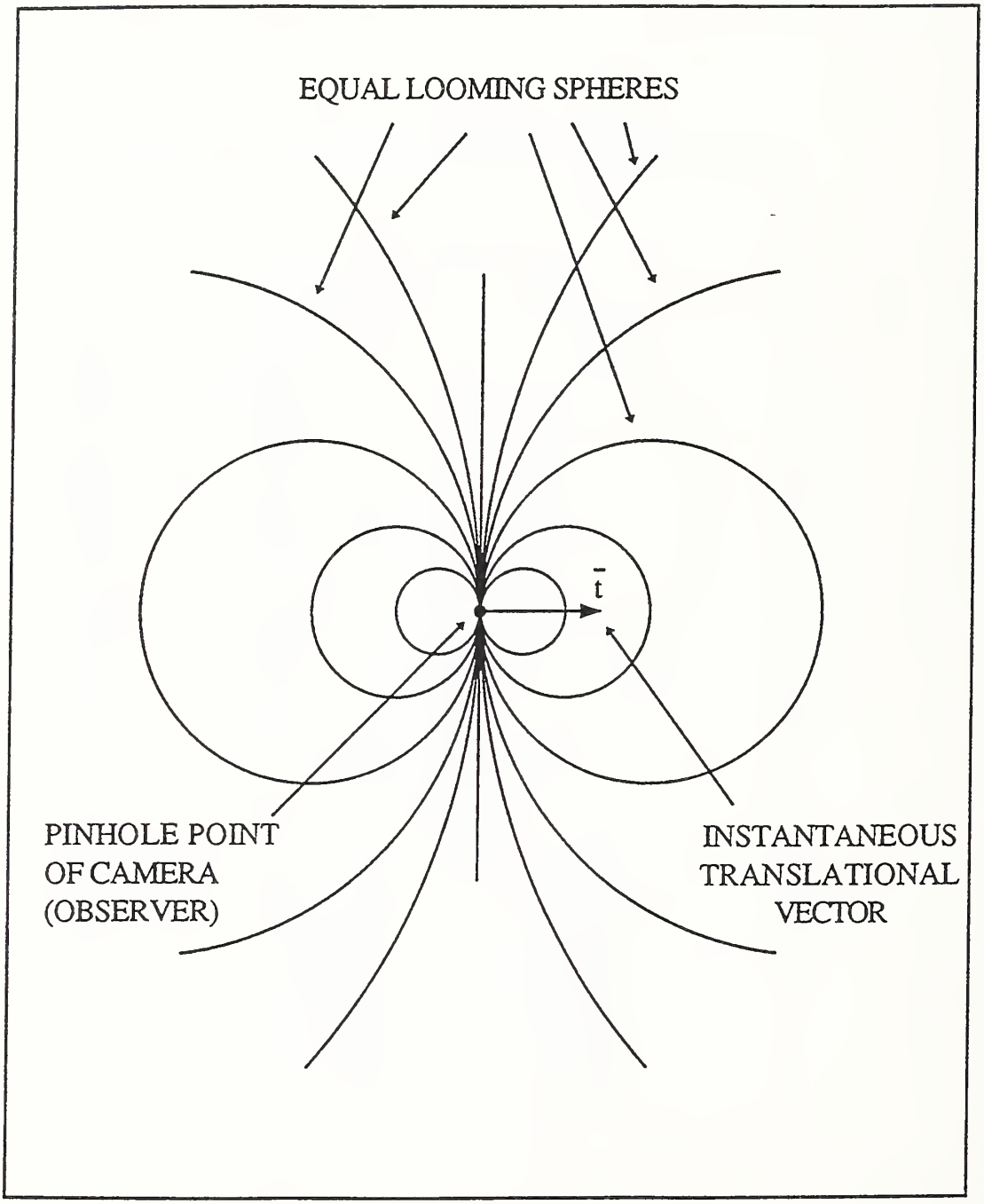

Figure 7: Equal Looming Spheres

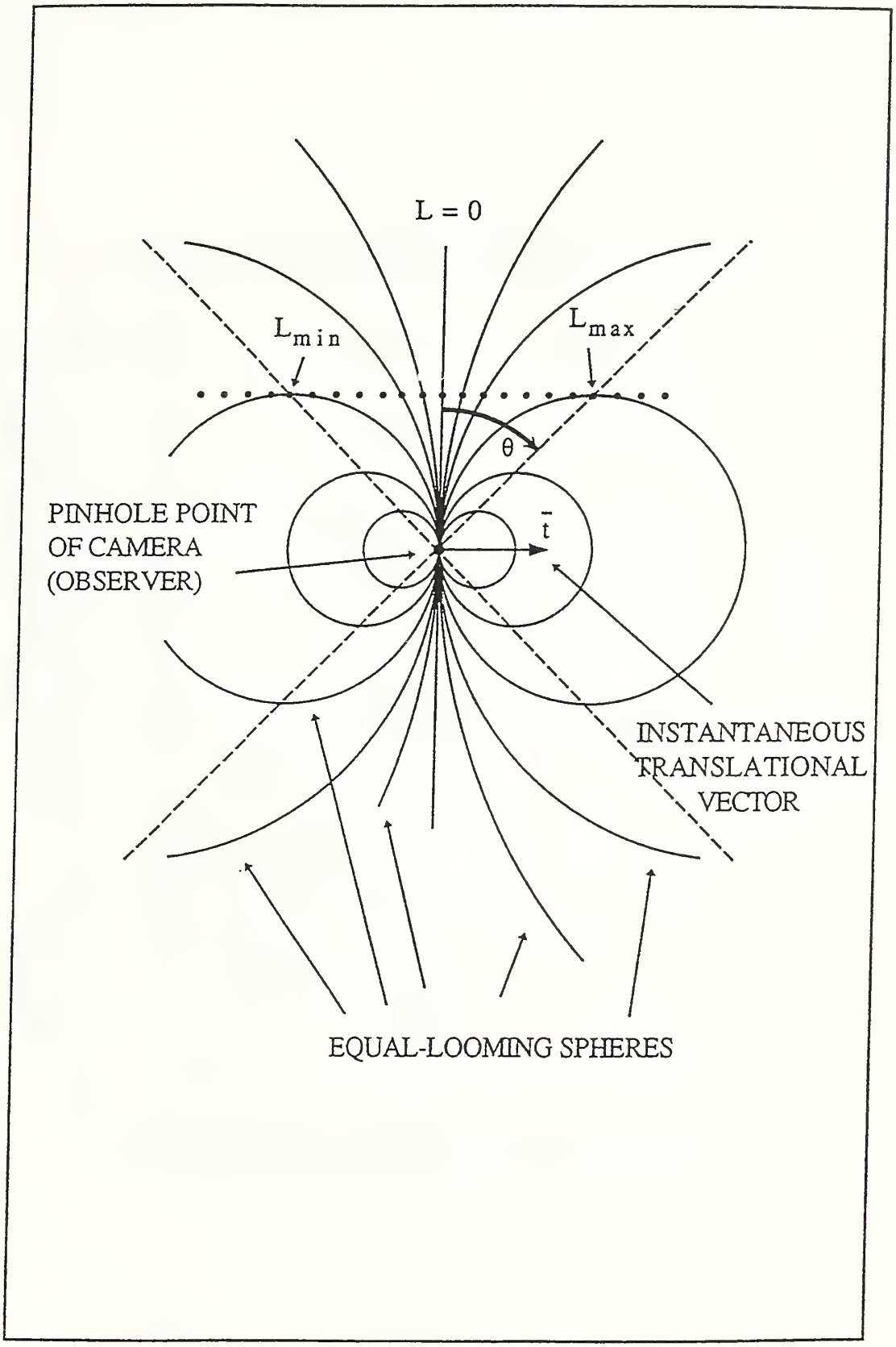

Figure 8: Equal Looming Spheres

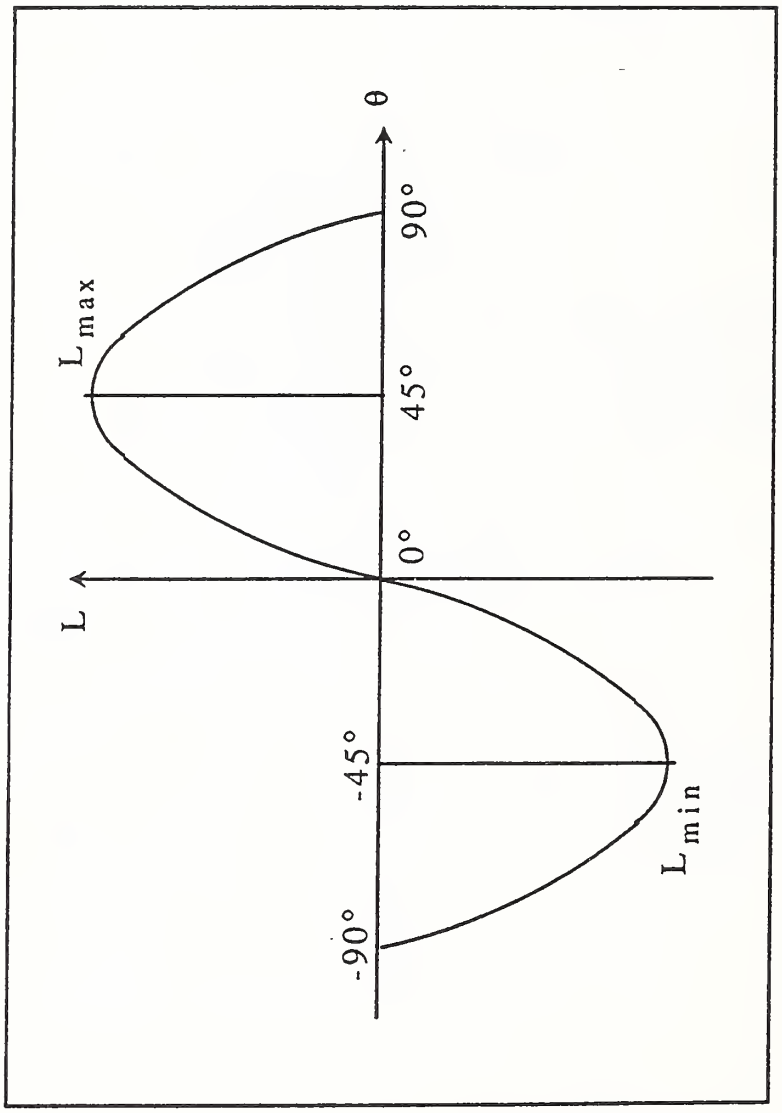

Figure 9: Looming for Different Angles

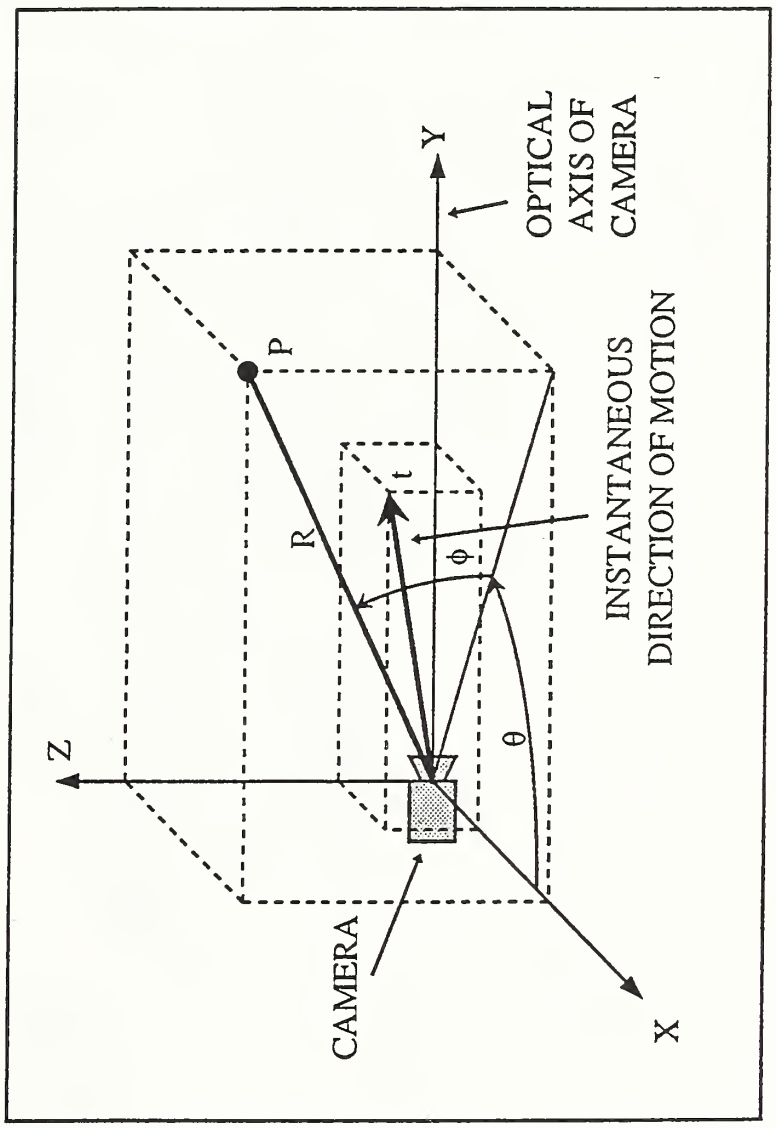

Figure 10: Camera Coordinate System

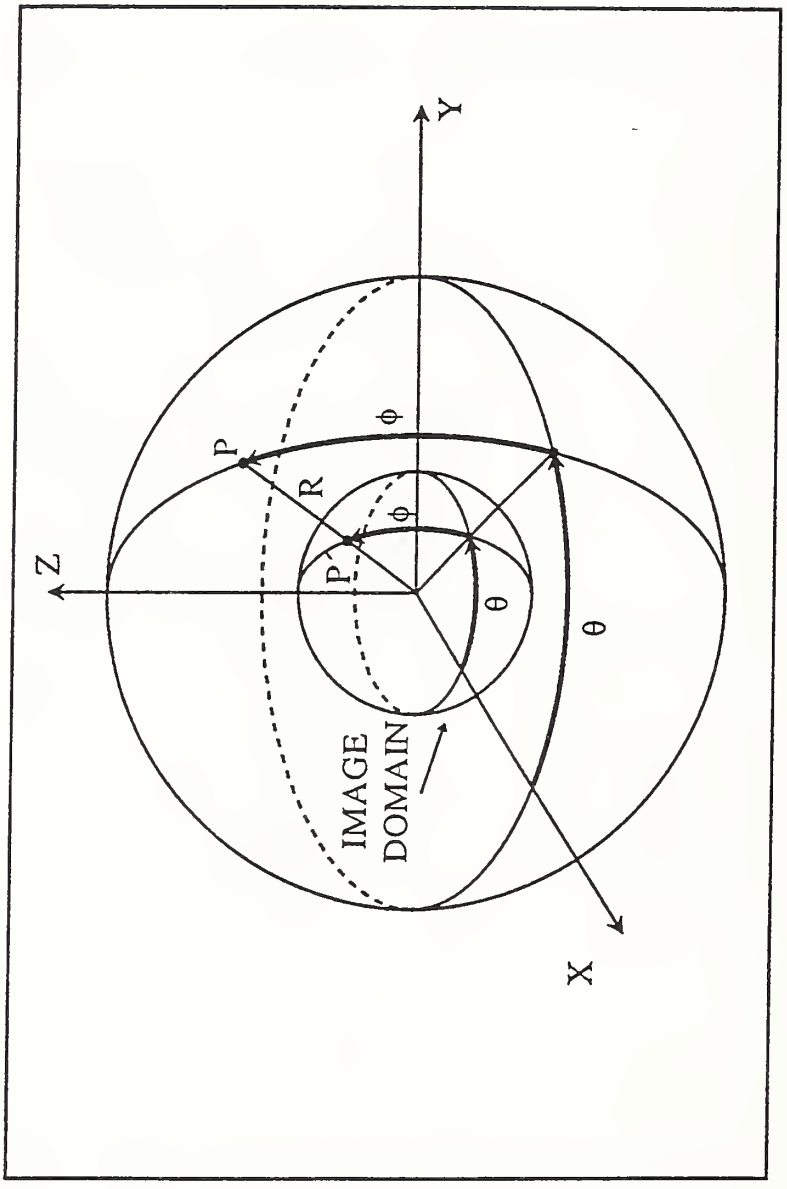

Figure 11: Image Domain

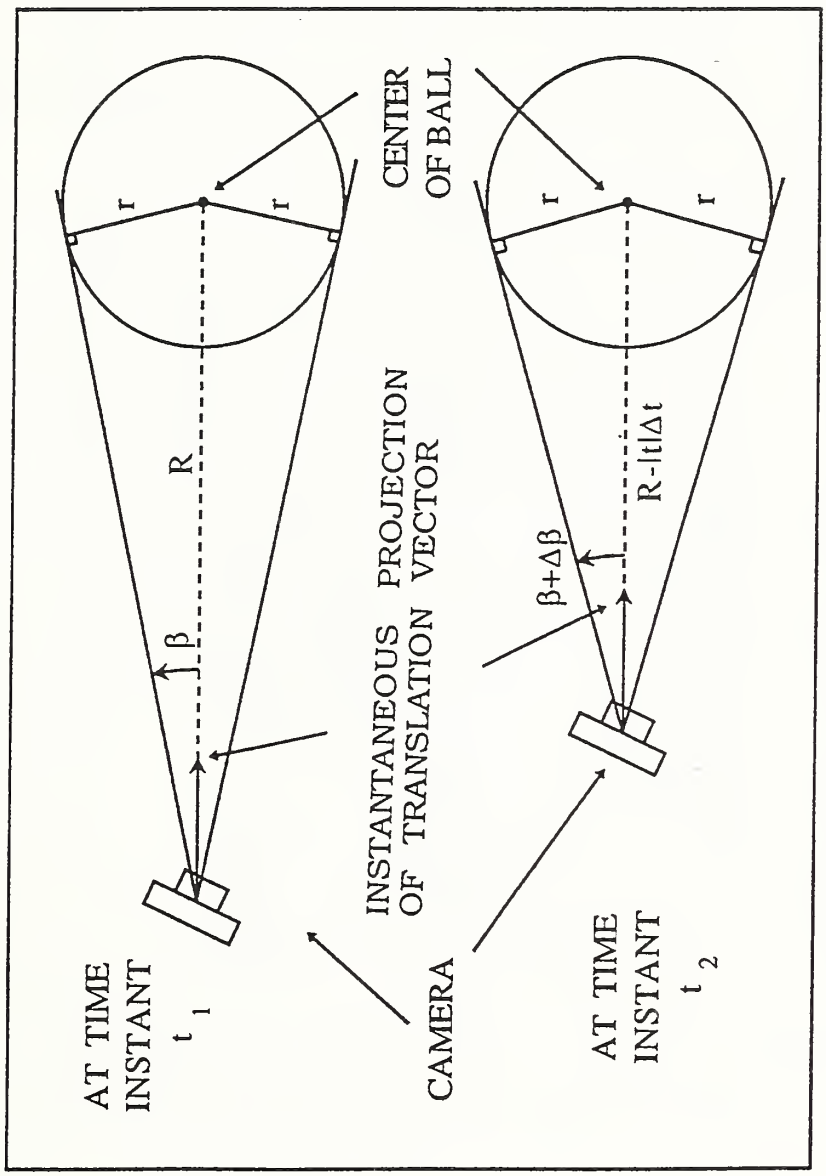

Figure 12: Looming Computations

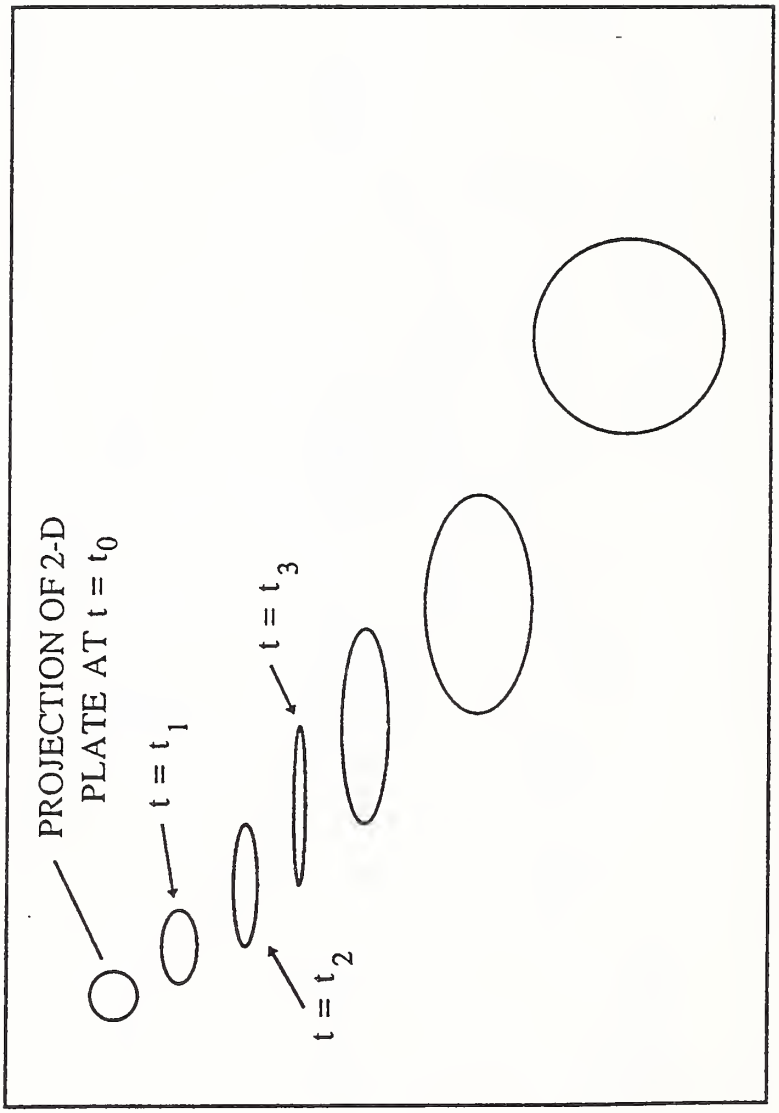

Figure 13: Projections of 2-D Plate

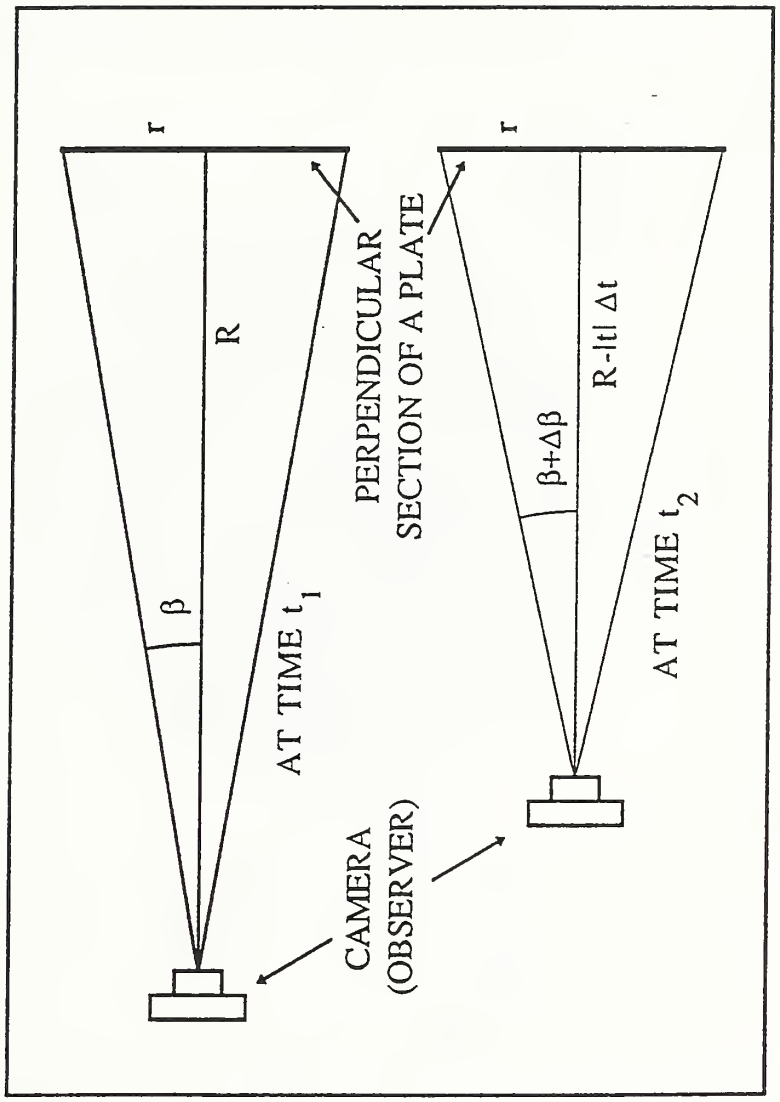

Figure 14: Moving 2-D Plate

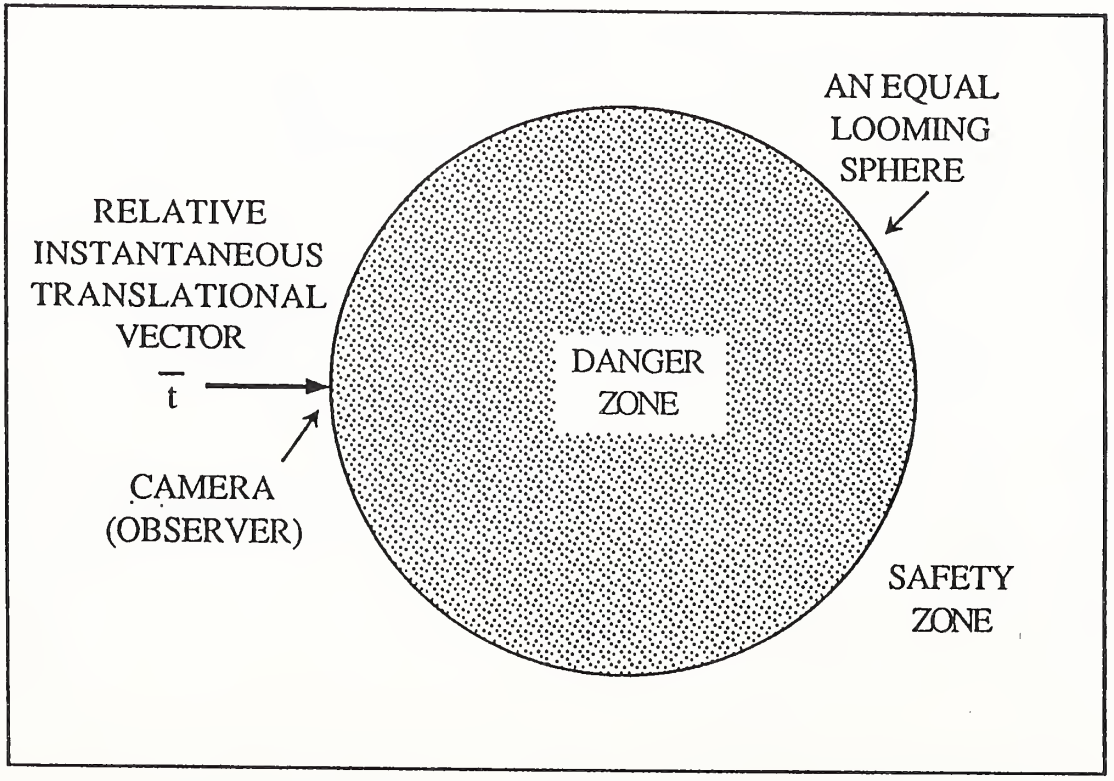

Figure 15(a): Qualitative Looming

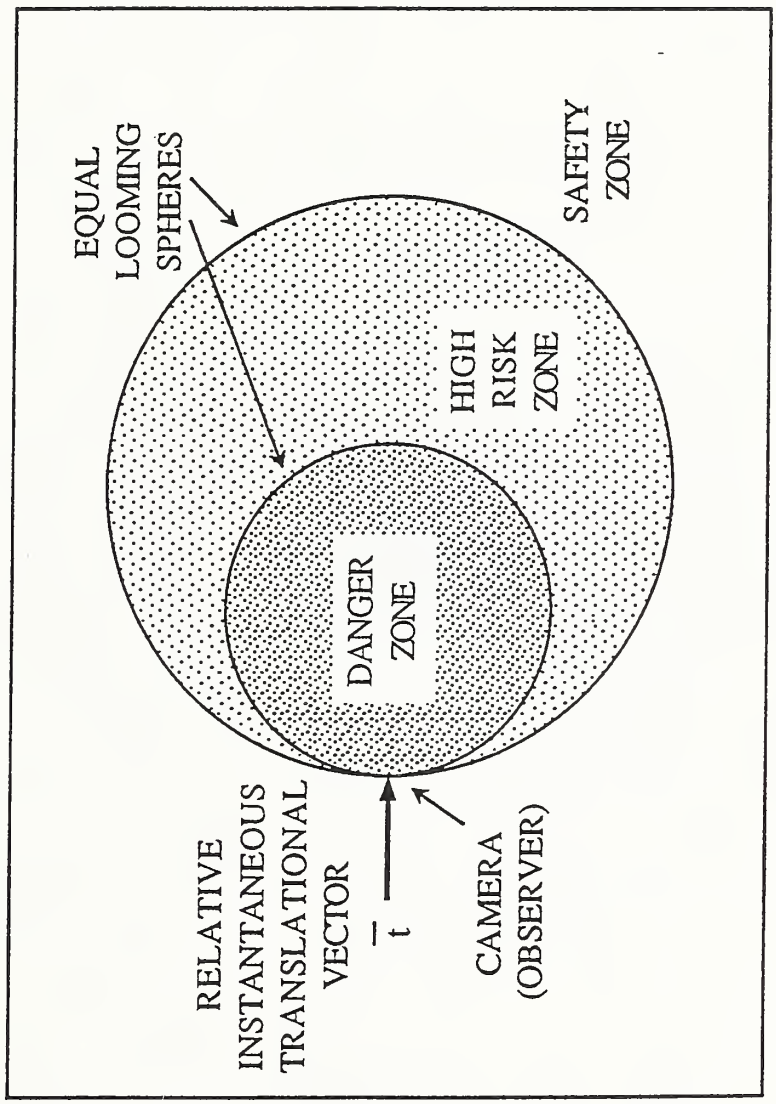

Figure 15(b): Qualitative Looming

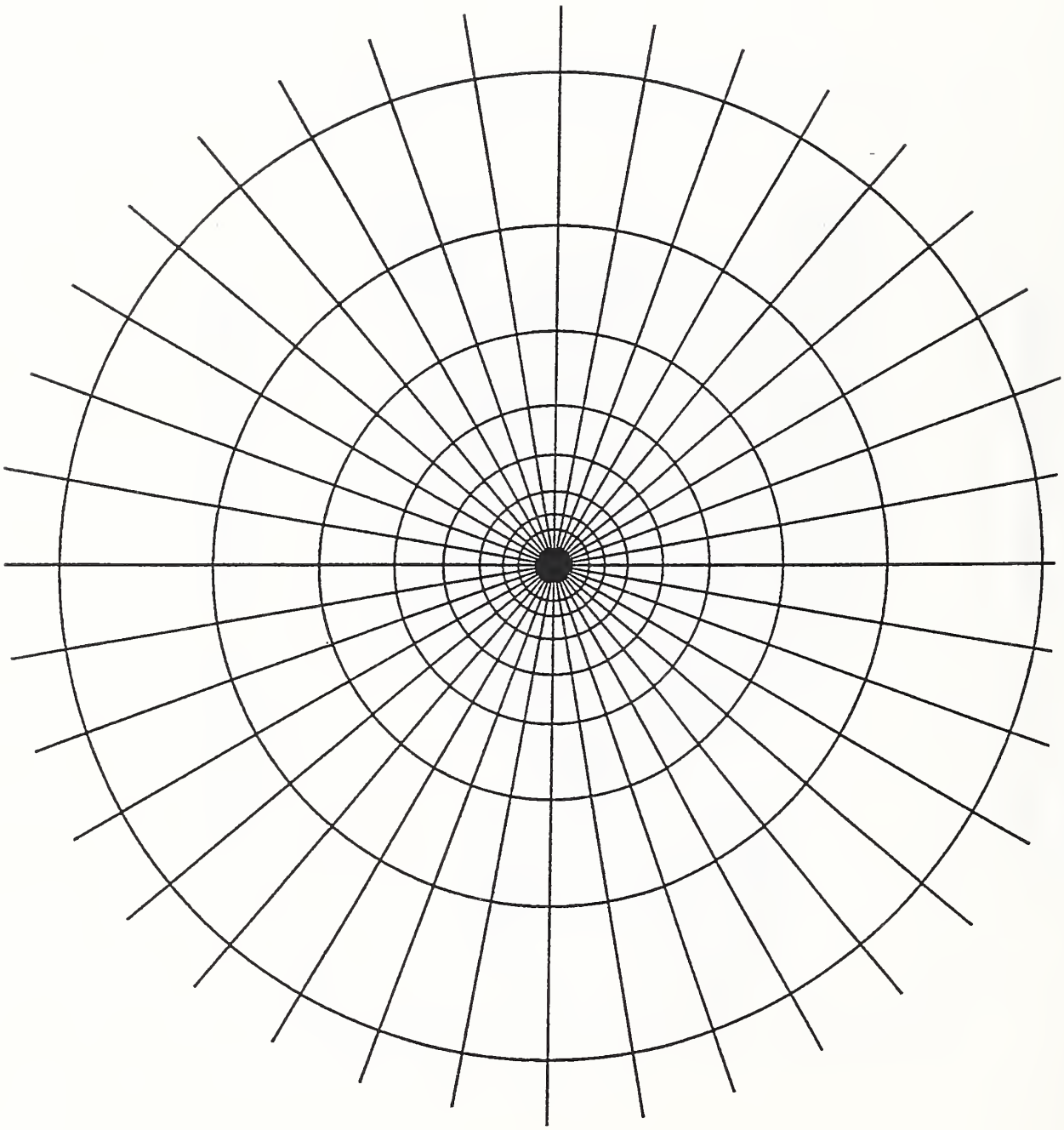

Figure 16: Logarithmic Retina

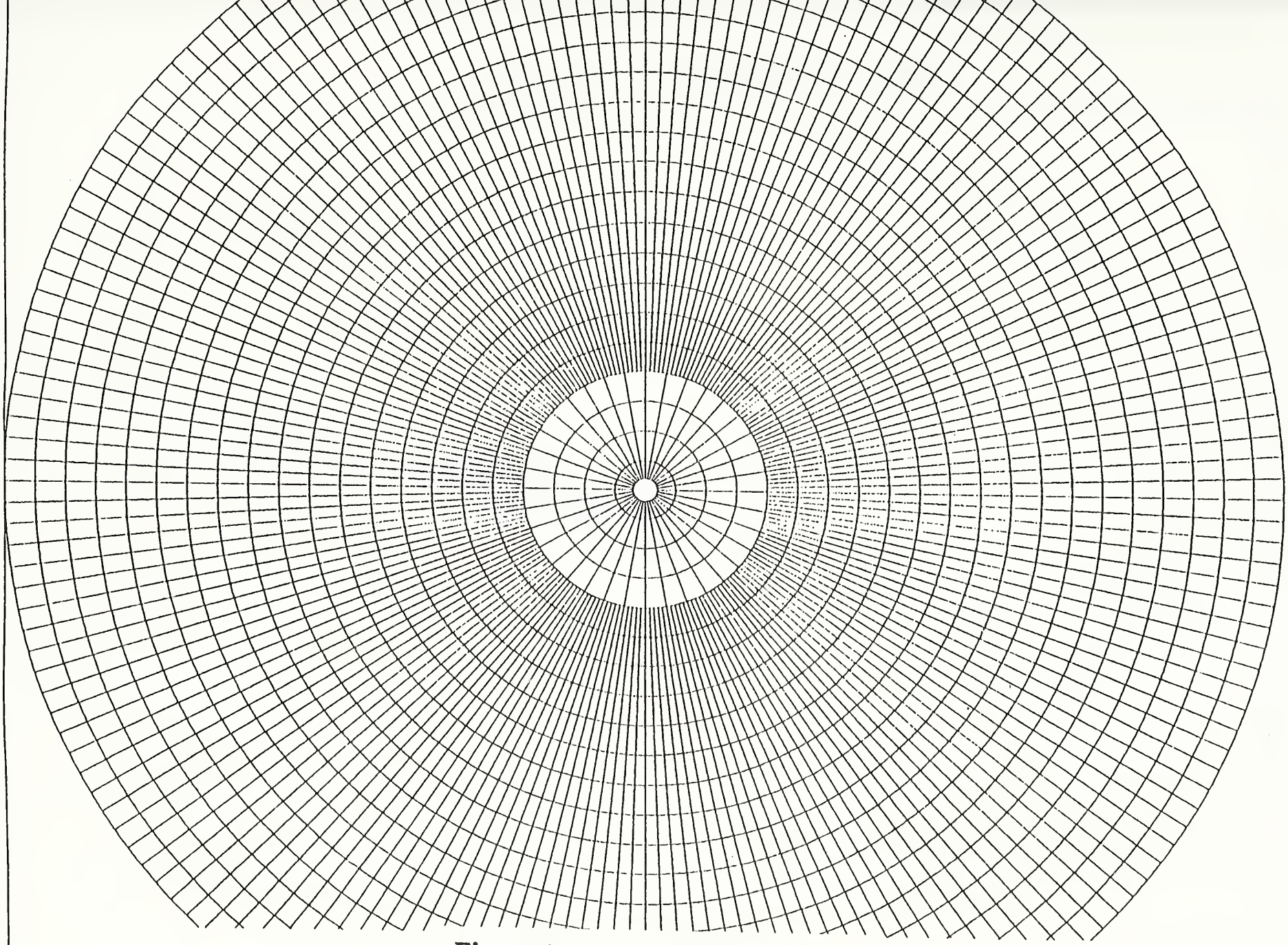

Figure 17: Linear Retina

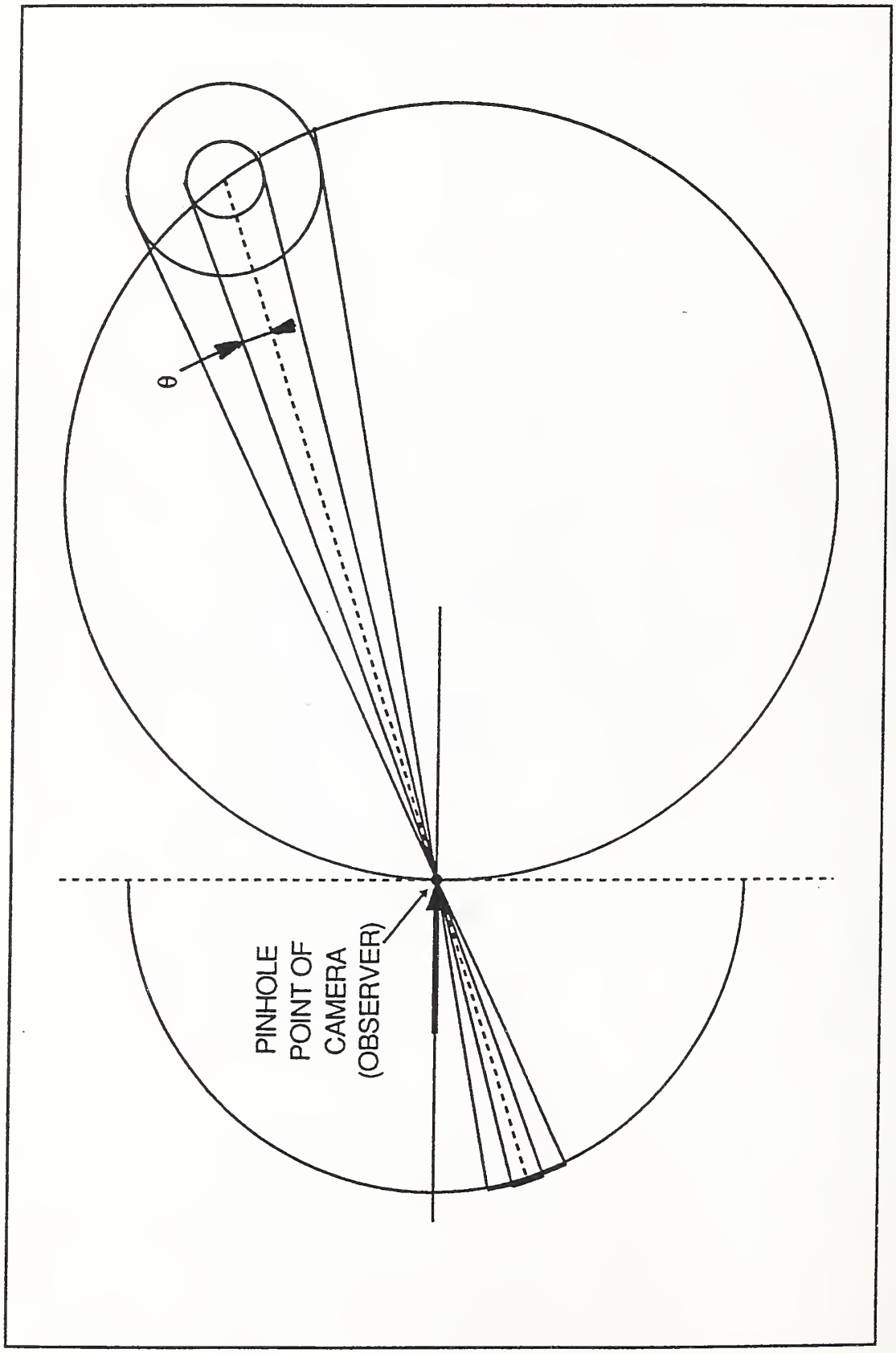

Figure 18: Looming and Fixation

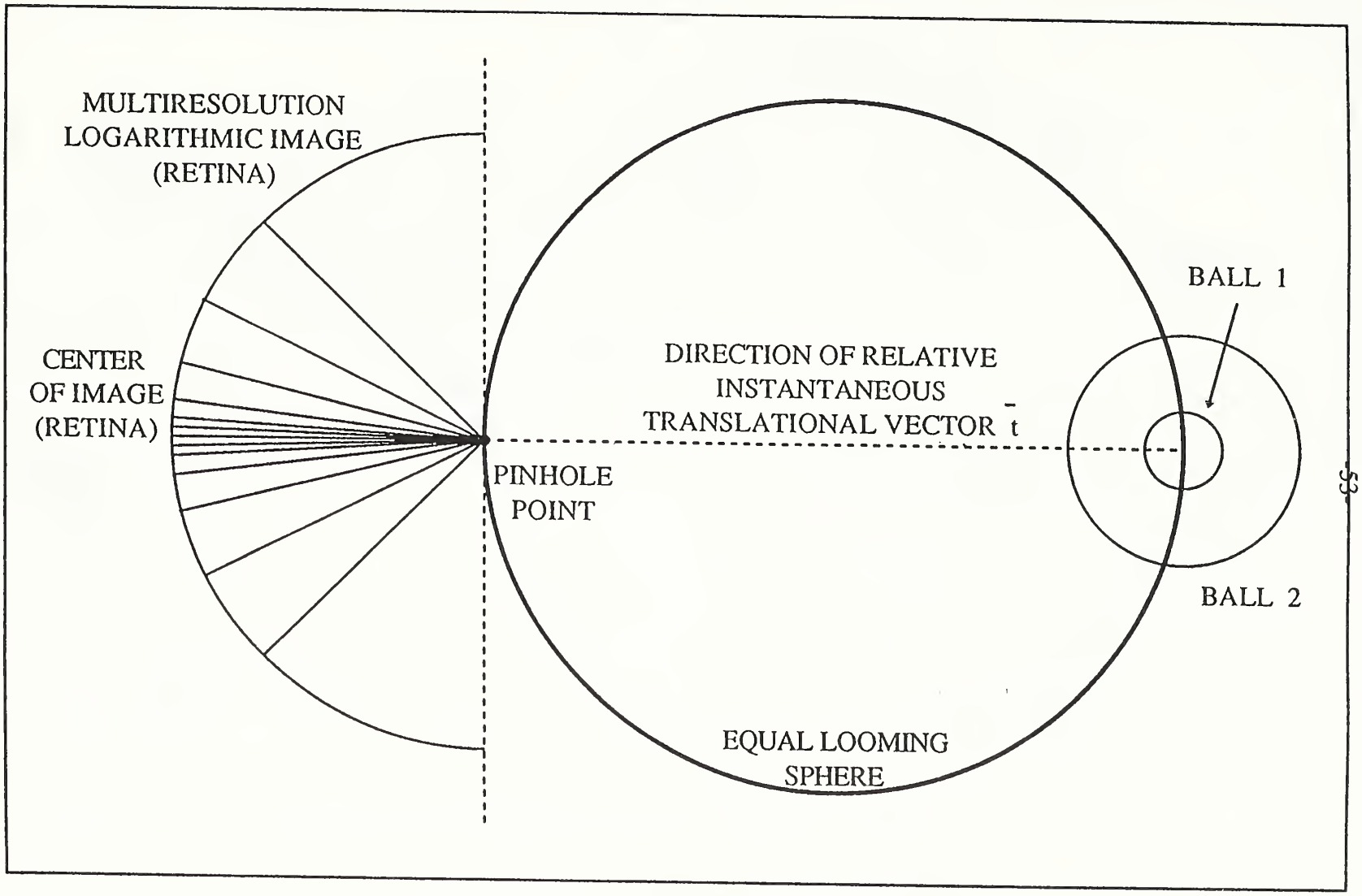

Figure 19: Looming and Fixation

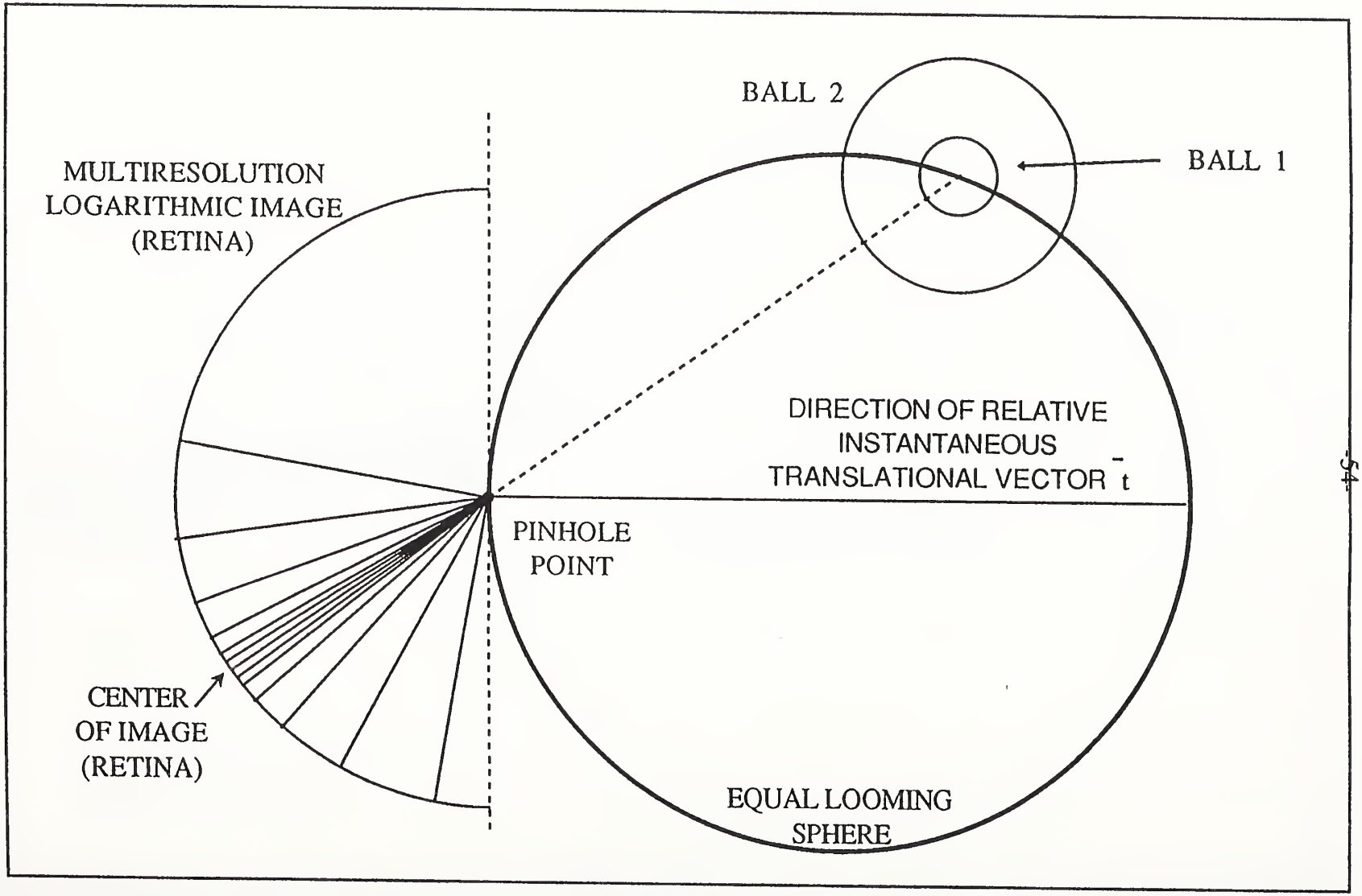

Figure 20: Looming and Fixation

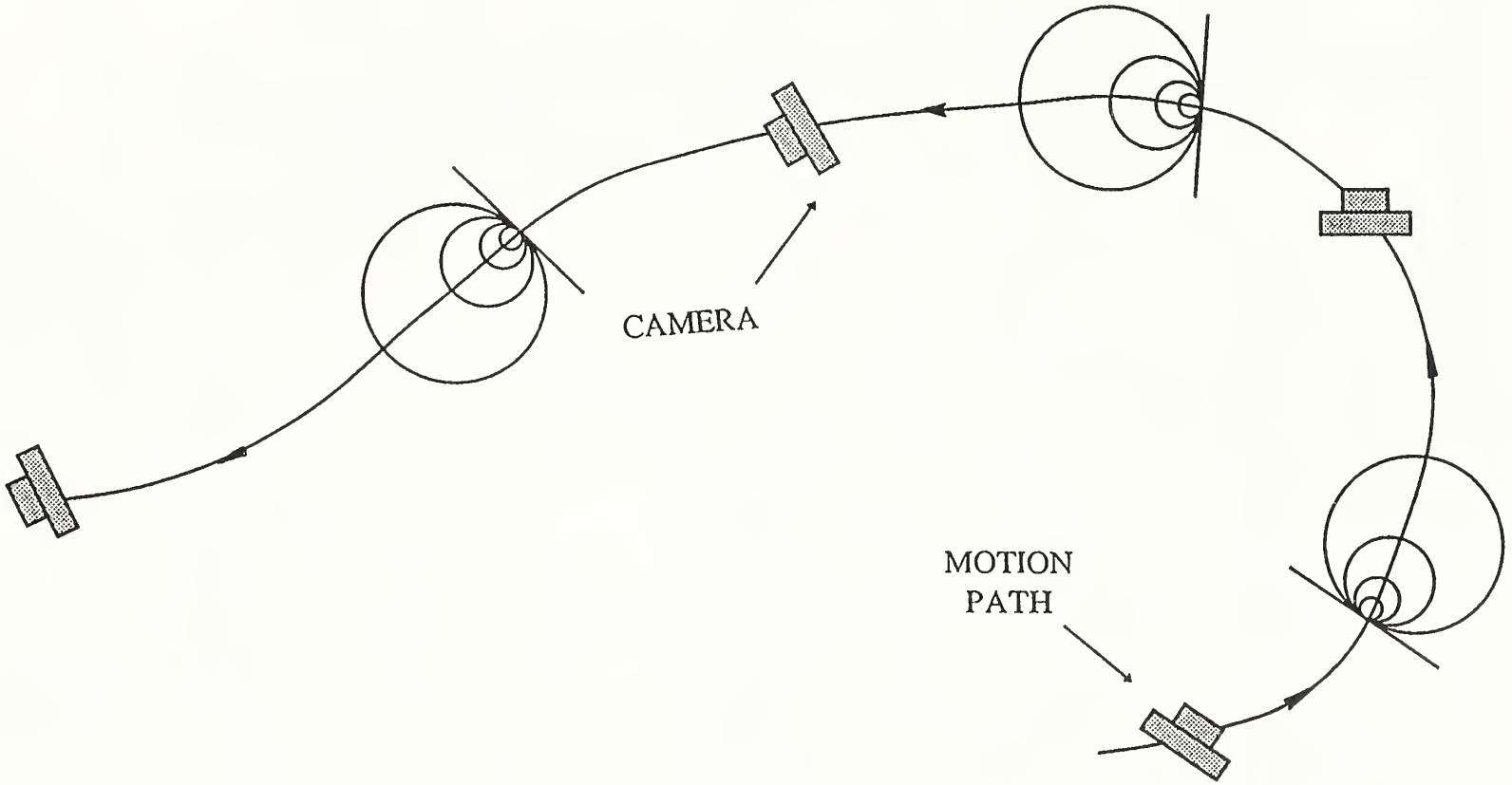

Figure 21: **Equal Looming Spheres as a Function of Time**

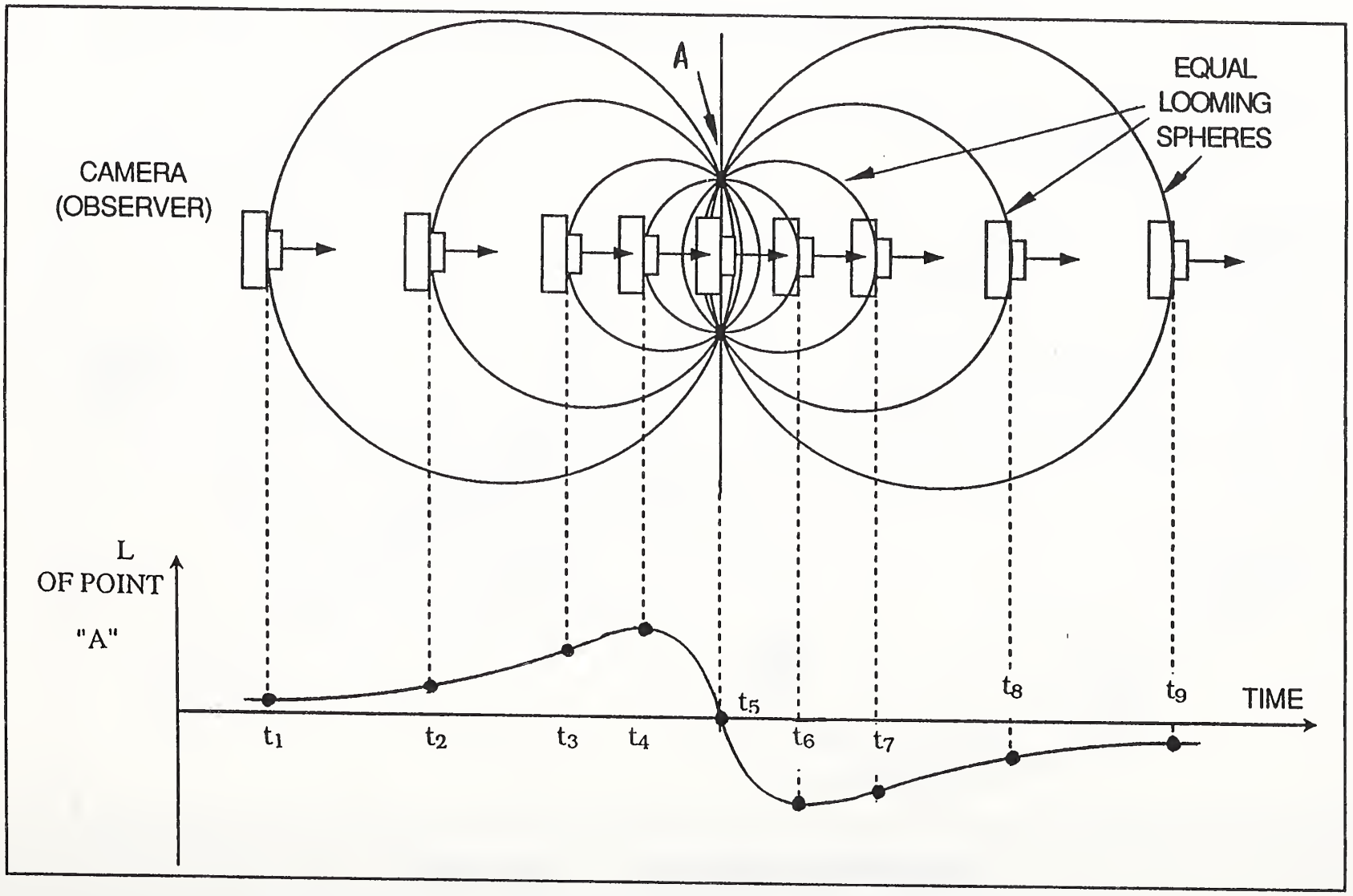

Figure 22: **Looming of Point "A" as a Function of Time**

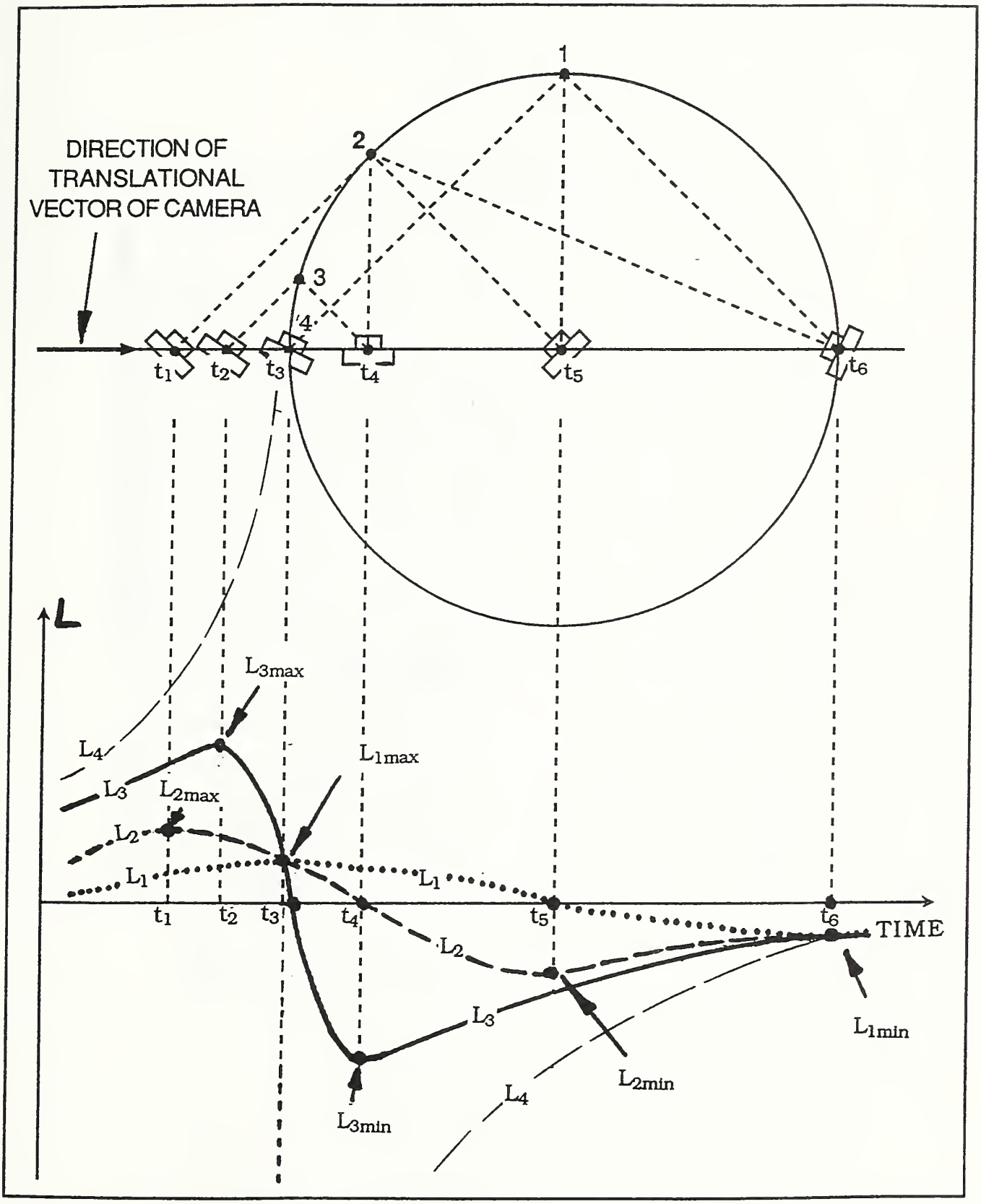

Figure 23: Looming as a Function of Time

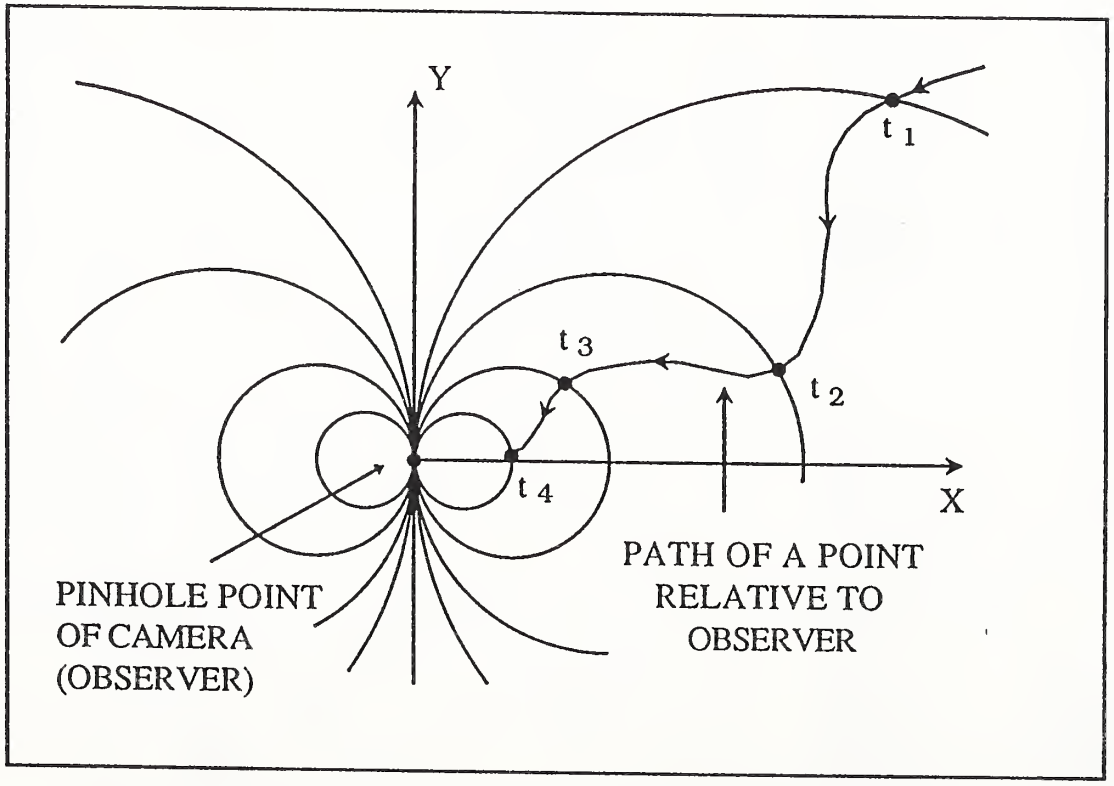

Figure 24(a): Keeping Constant Looming Value

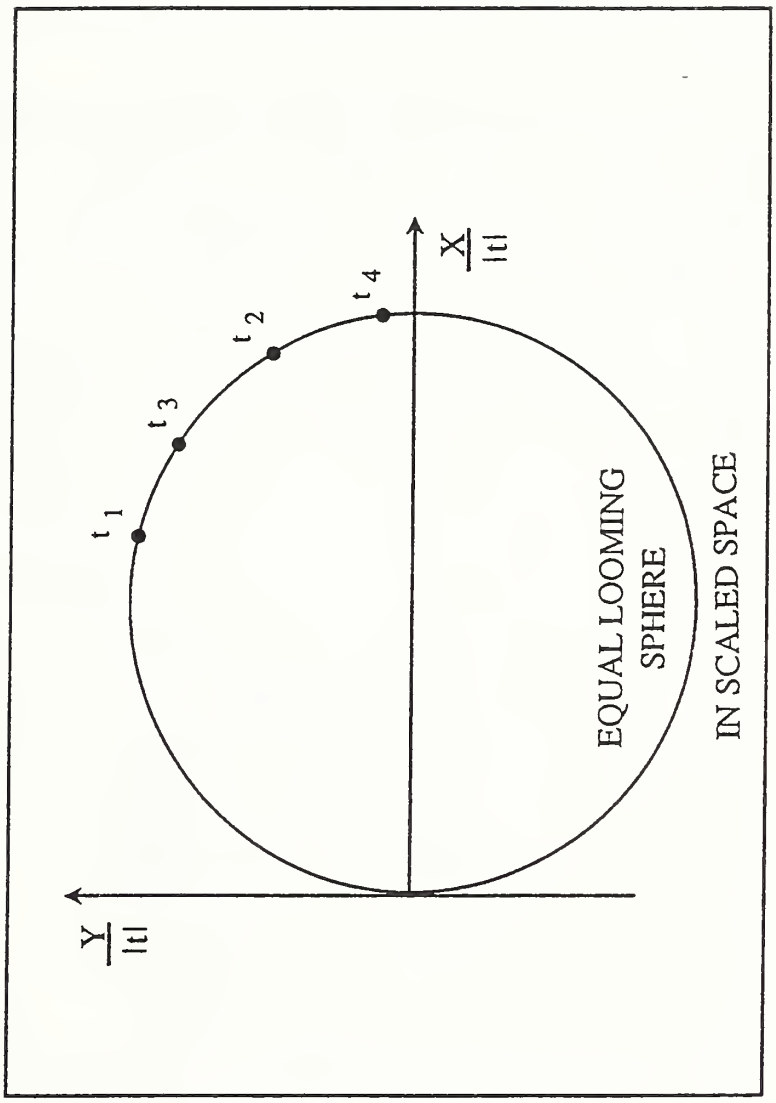

Figure 24(b): Keeping Constant Looming Value

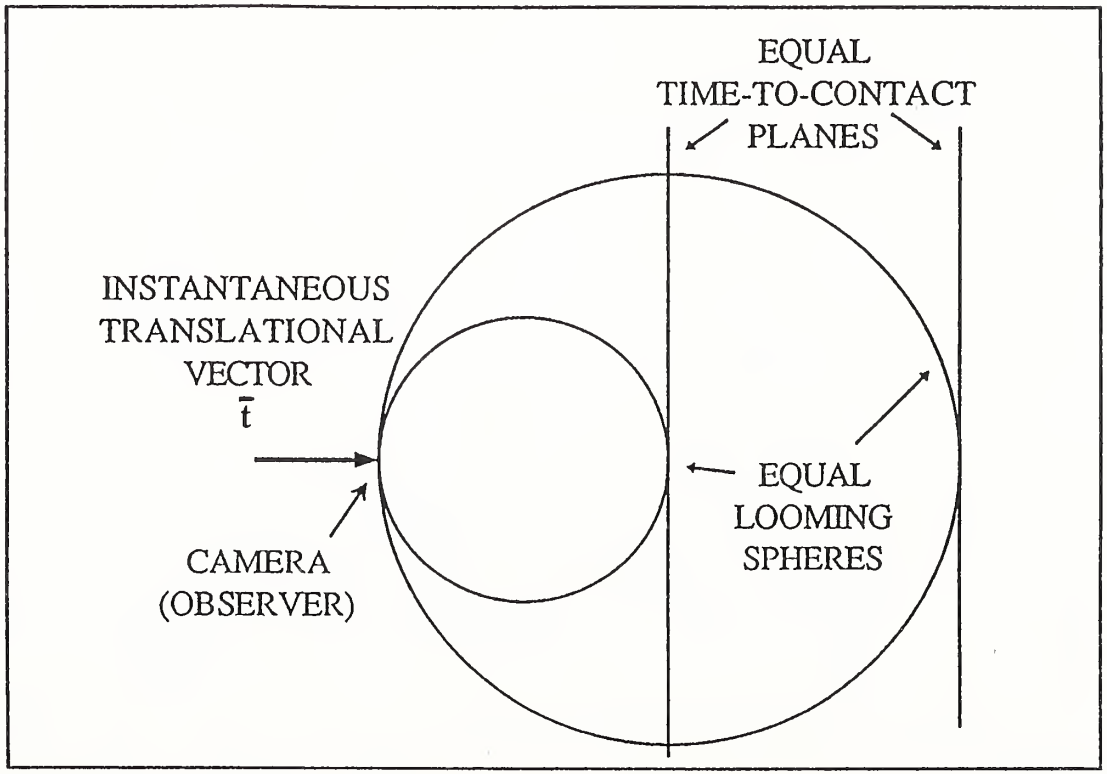

Figure 25: Looming and Time-to-Contact

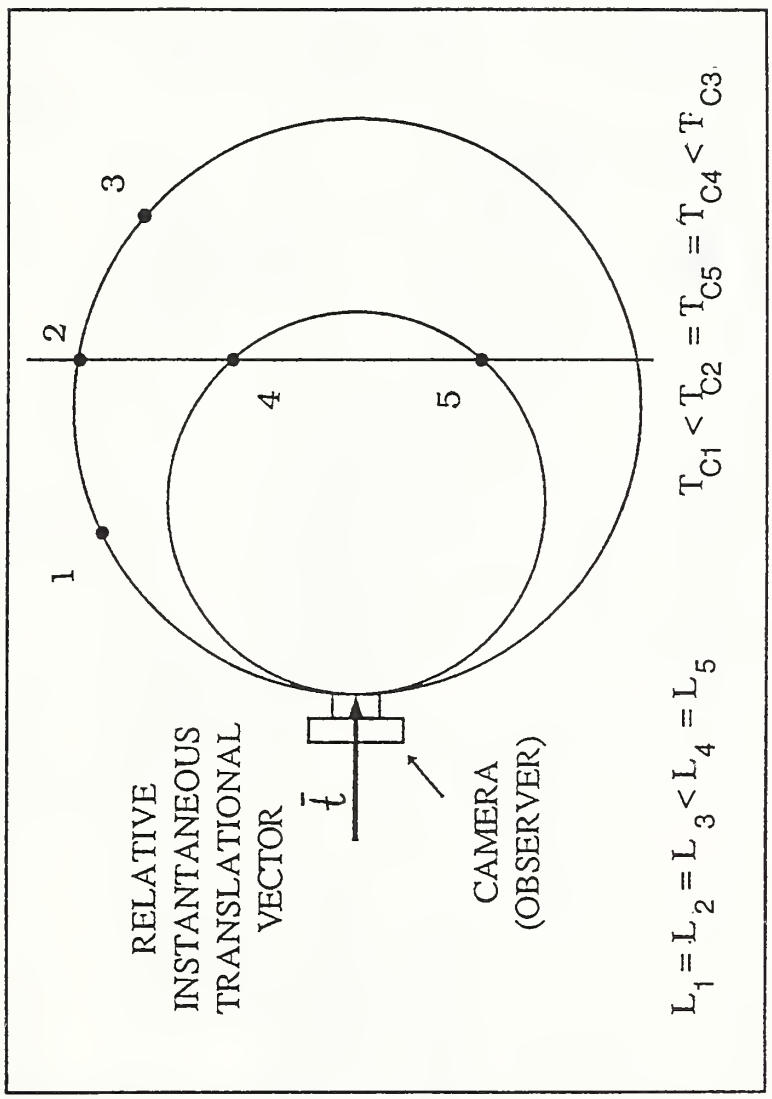

Figure 26: Looming and Time-to-Contact

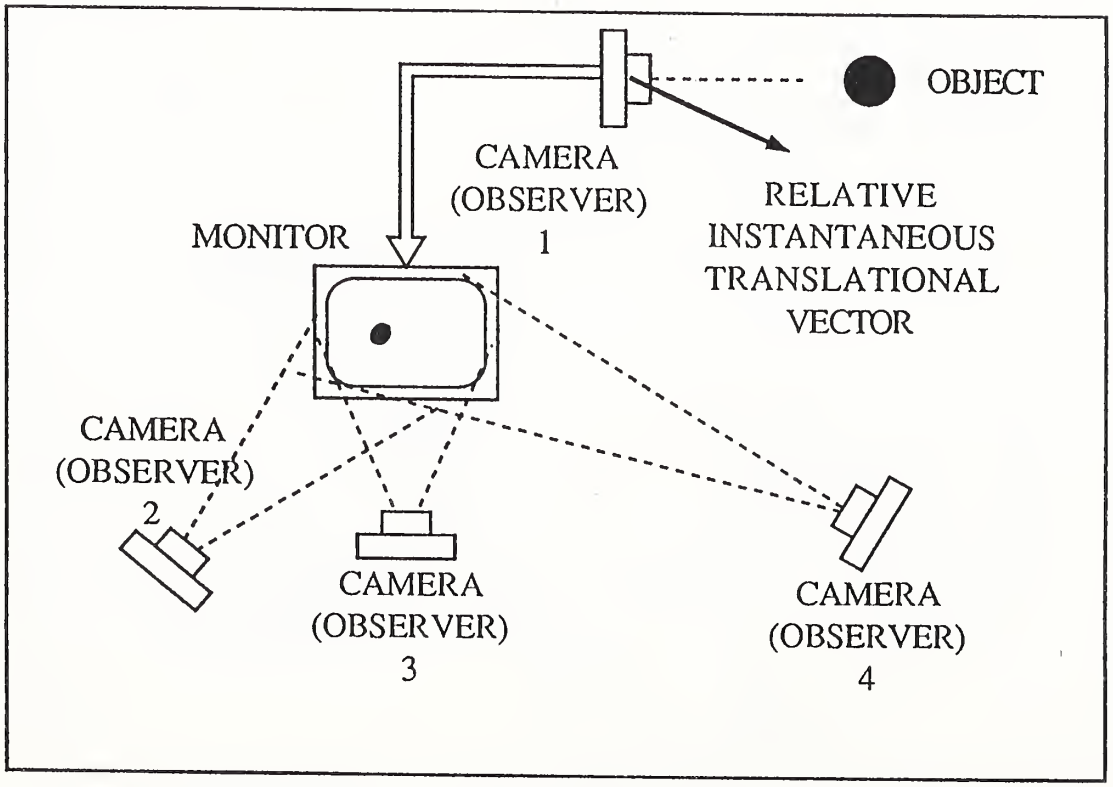

Figure 27: Observing a Screen

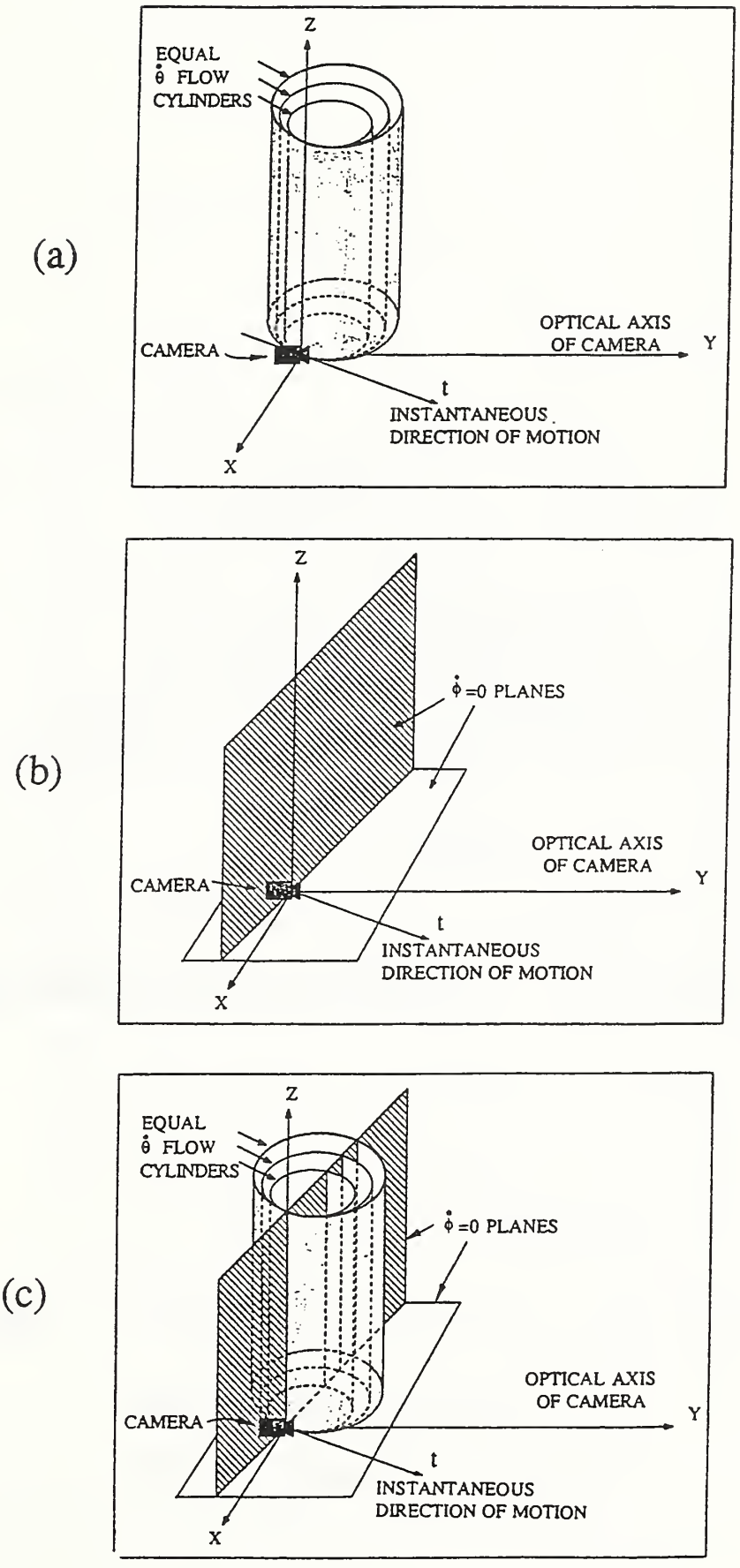

Figure 28: (a) Equal $\dot{\theta}$ Cylinders
(b) Equal $\dot{\phi}$=0 Planes
(c) IntersectionBbetween $\dot{\theta}$ Cylinders and $\dot{\phi}$=0 Planes

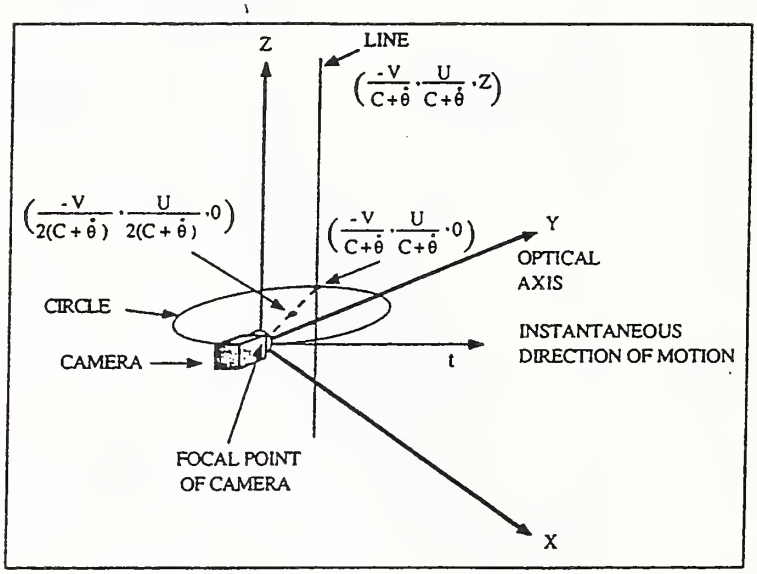

Figure 29: Pictorial Description of Solutions

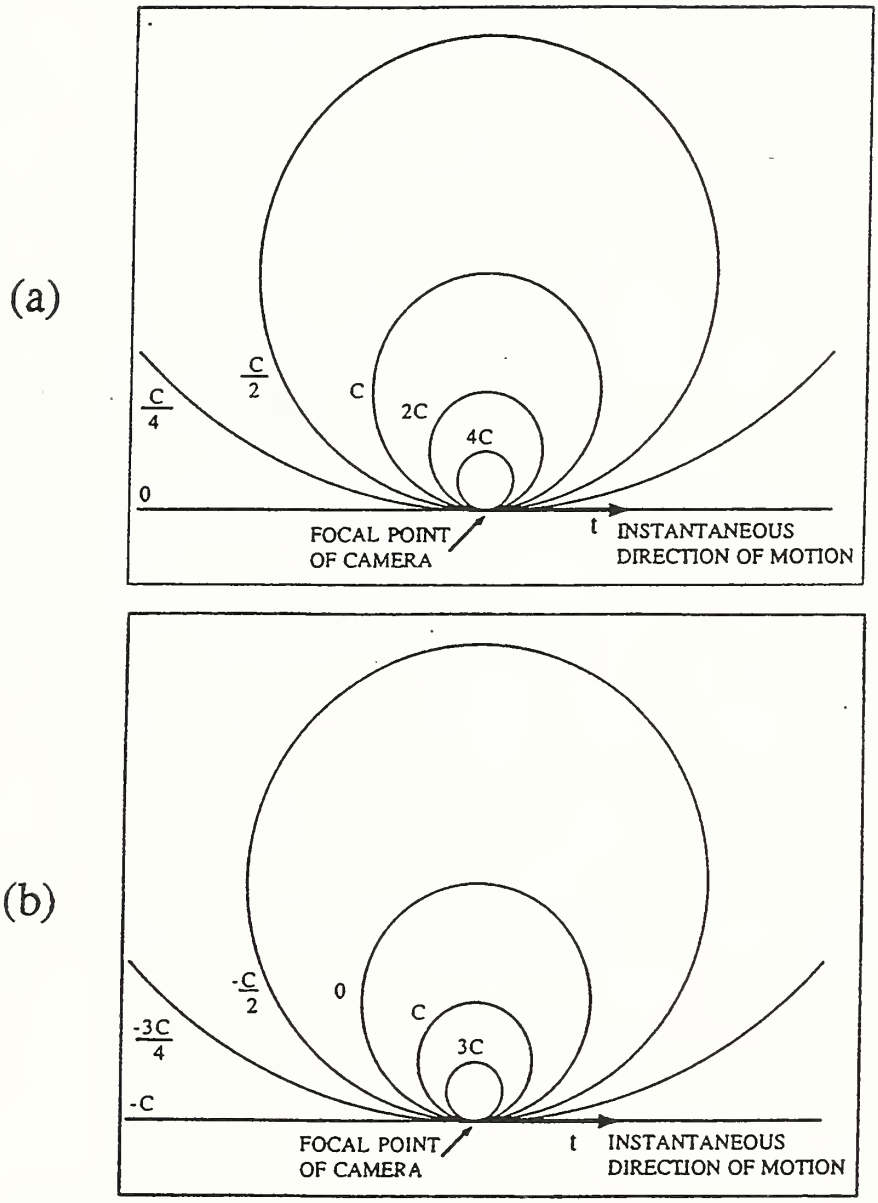

Figure 30: **Optical Flow Values Due to Camera Motion**

(a) For Translation Only

(b) For Translation and Rotation

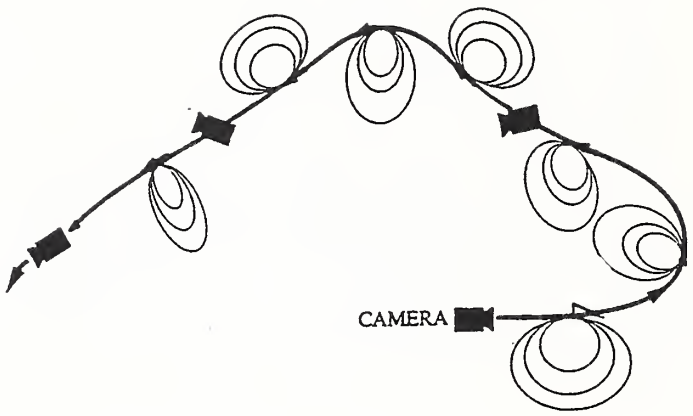

Figure 31: EFC'S as a Function of Time

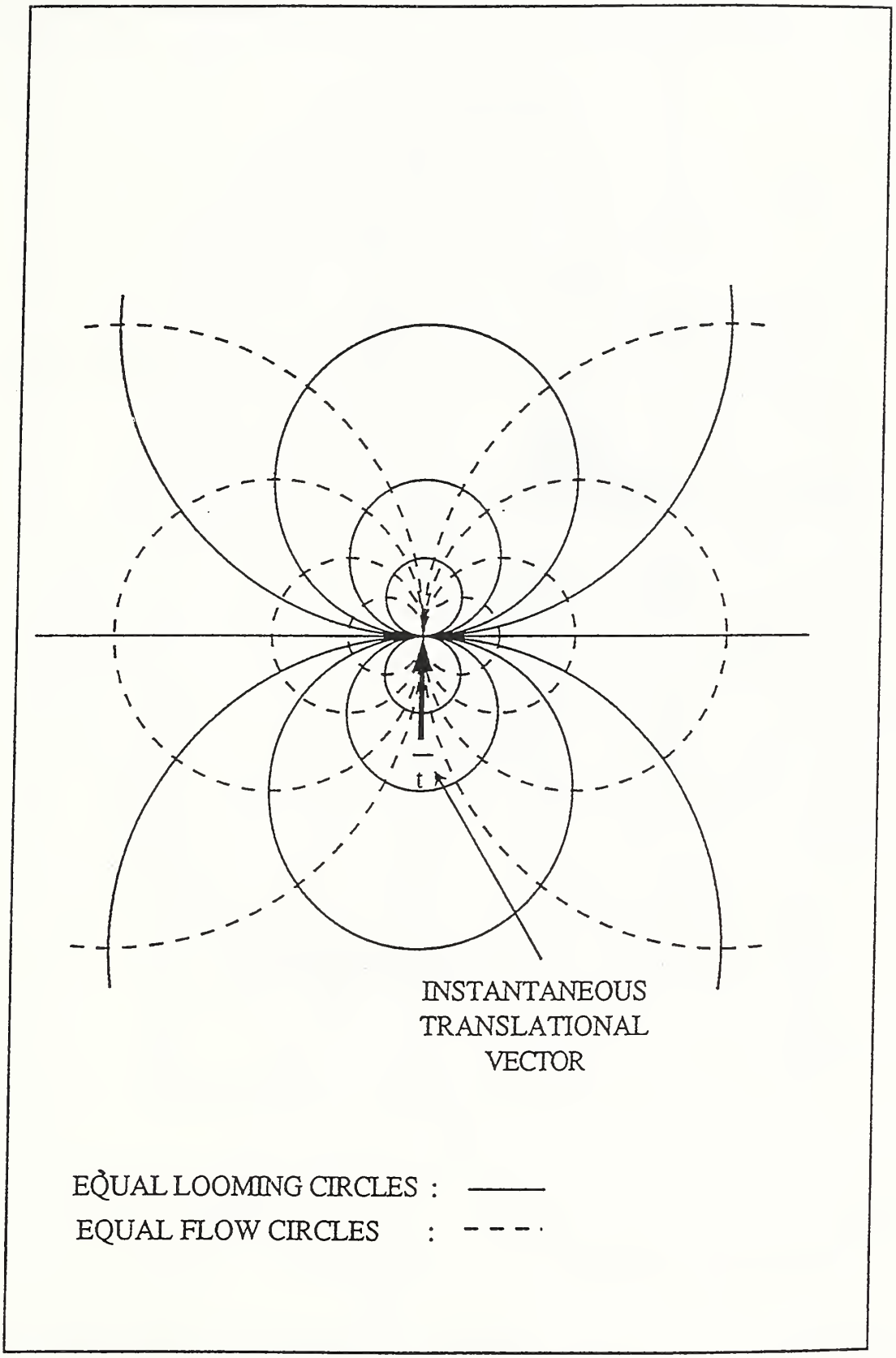

Figure 32: EFC'S-Looming Mapping

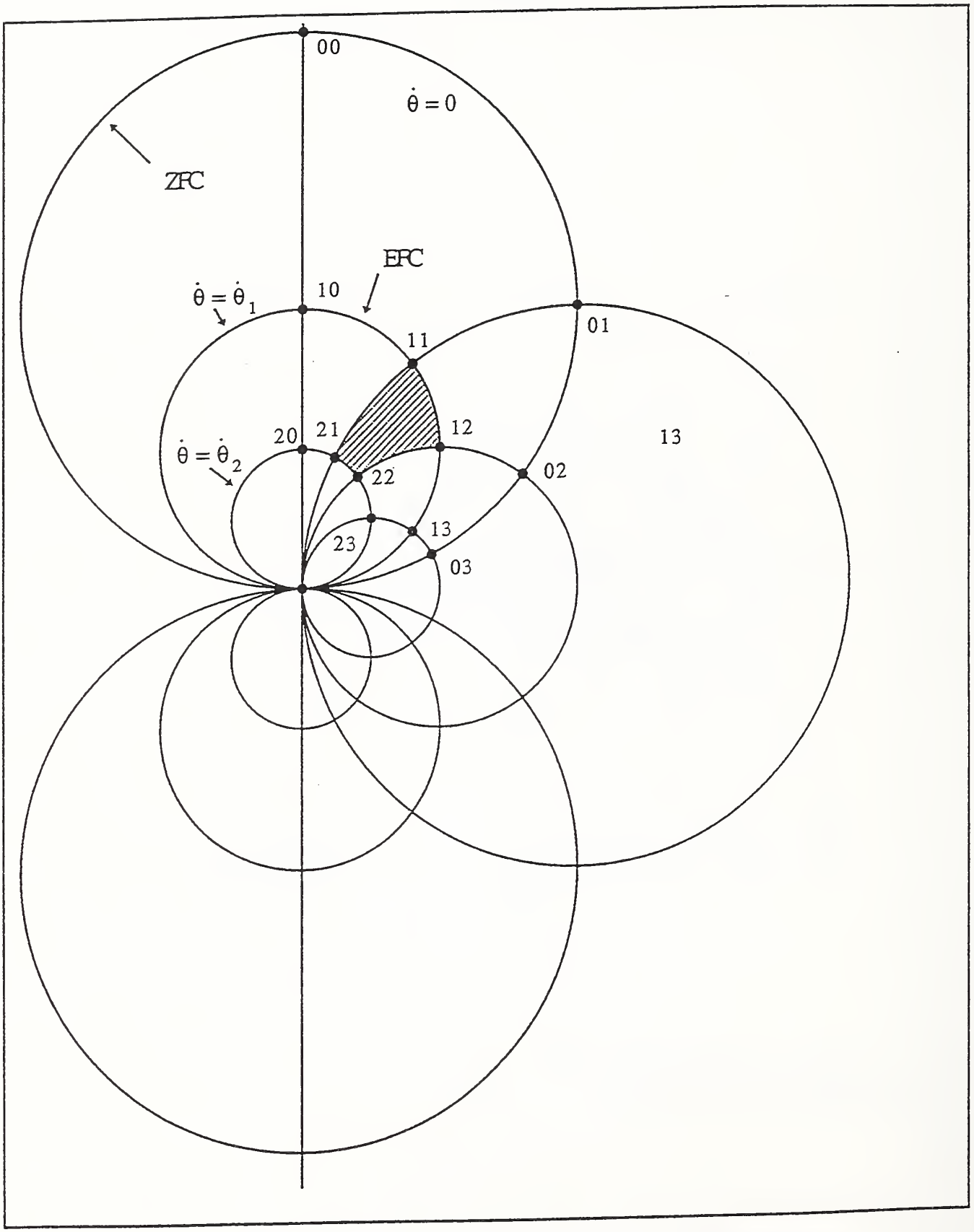

Figure 33: EFC'S-Looming Mapping

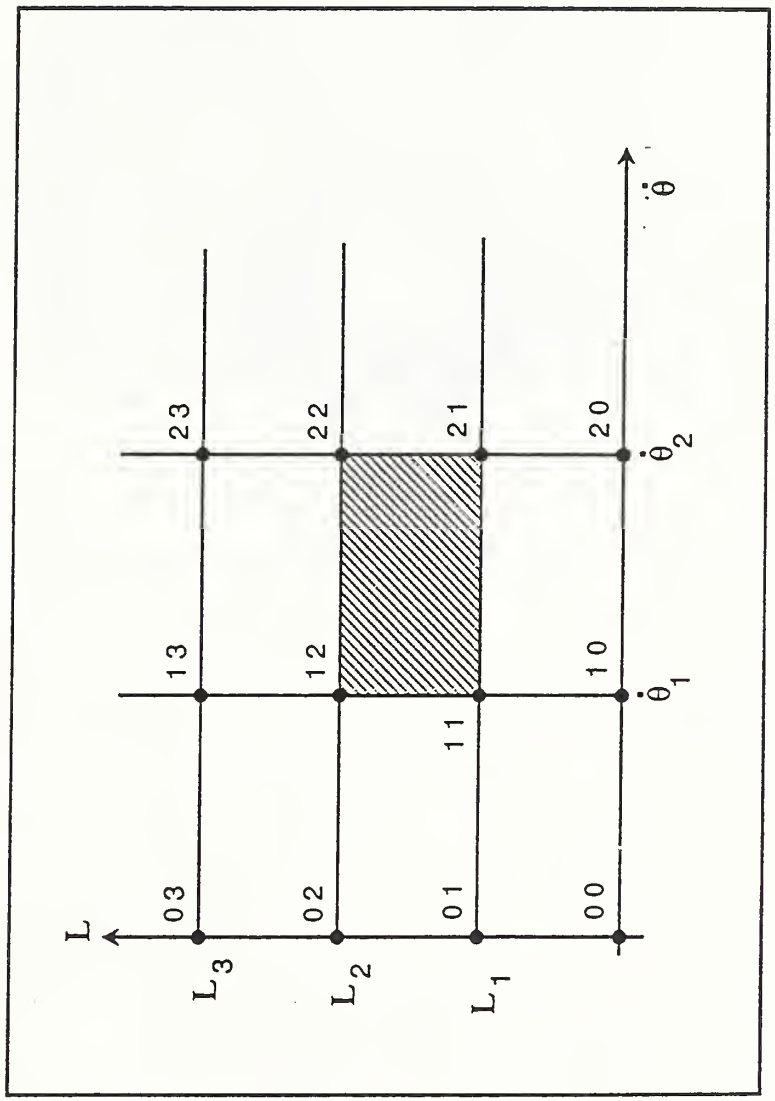

Figure 34: Isometric Mapping of Looming and EFC'S

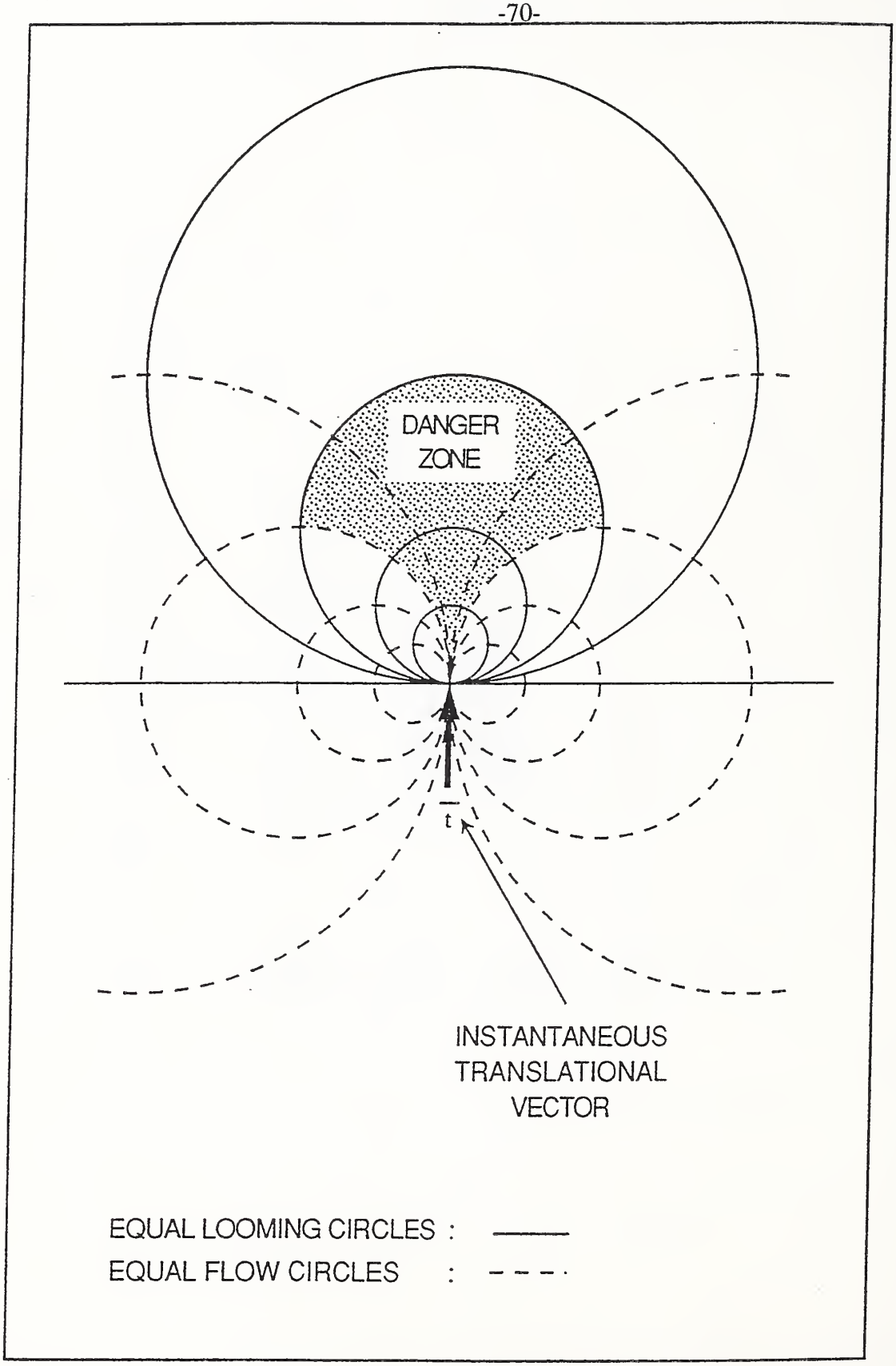

Figure 35: Defining Obstacles Using Looming

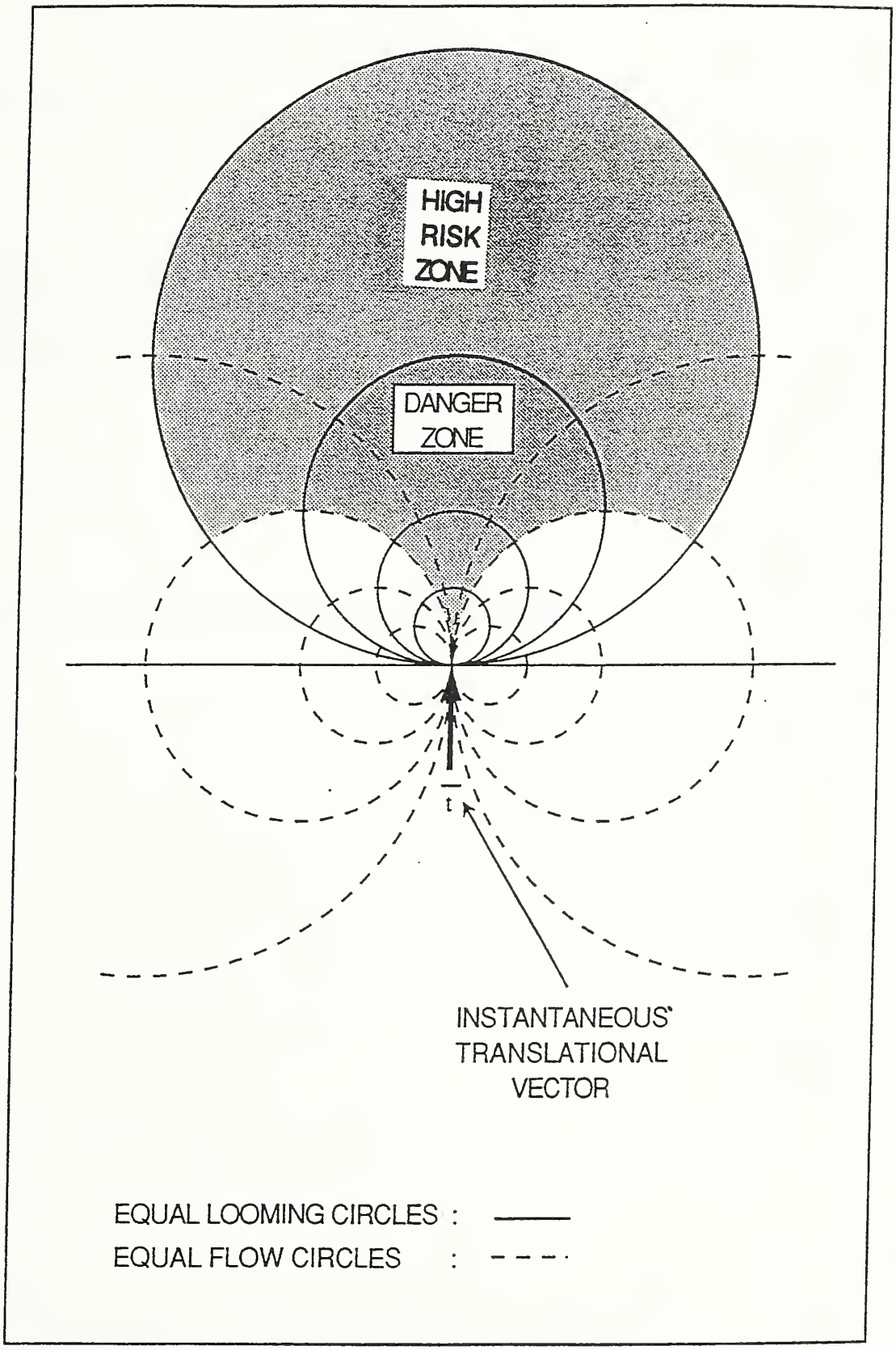

Figure 36: Defining Obstacles Using Looming

NIST-114A
(REV. 3-90)

U.S. DEPARTMENT OF COMMERCE
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY

BIBLIOGRAPHIC DATA SHEET

1. PUBLICATION OR REPORT NUMBER
NISTIR 4808

2. PERFORMING ORGANIZATION REPORT NUMBER

3. PUBLICATION DATE
APRIL 1992

4. TITLE AND SUBTITLE

A Quantitative Approach to Looming

5. AUTHOR(S)

Daniel Raviv

6. PERFORMING ORGANIZATION (IF JOINT OR OTHER THAN NIST, SEE INSTRUCTIONS)

U.S. DEPARTMENT OF COMMERCE
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
GAITHERSBURG, MD 20899

7. CONTRACT/GRANT NUMBER

8. TYPE OF REPORT AND PERIOD COVERED

9. SPONSORING ORGANIZATION NAME AND COMPLETE ADDRESS (STREET, CITY, STATE, ZIP)

10. SUPPLEMENTARY NOTES

11. ABSTRACT (A 200-WORD OR LESS FACTUAL SUMMARY OF MOST SIGNIFICANT INFORMATION. IF DOCUMENT INCLUDES A SIGNIFICANT BIBLIOGRAPHY OR LITERATURE SURVEY, MENTION IT HERE.)

ABSTRACT

The visual looming effect has been shown to be very important when action is taking place. Most existing work on "looming" is qualitative or limited-quantitative. In this paper we take a quantitative approach to understanding "looming". We define looming mathematically, show geometrical properties of objects that produce the same value of looming, explain how to measure looming in the general case of motion, how a multiresolution logarithmic retina simplifies the measurement of looming, and how the results can be combined with previous work on visual fields. We suggest a new representation of space based on looming and the so called Equal Flow Circles (EFCs).

12. KEY WORDS (6 TO 12 ENTRIES; ALPHABETICAL ORDER; CAPITALIZE ONLY PROPER NAMES; AND SEPARATE KEY WORDS BY SEMICOLONS)

Active Vision; Vision-Based Behavior; Looming; Vision-Based Mobility

13. AVAILABILITY

[X] UNLIMITED

[ ] FOR OFFICIAL DISTRIBUTION. DO NOT RELEASE TO NATIONAL TECHNICAL INFORMATION SERVICE (NTIS).

[ ] ORDER FROM SUPERINTENDENT OF DOCUMENTS, U.S. GOVERNMENT PRINTING OFFICE, WASHINGTON, DC 20402.

[ ] ORDER FROM NATIONAL TECHNICAL INFORMATION SERVICE (NTIS), SPRINGFIELD, VA 22161.

14. NUMBER OF PRINTED PAGES
75

15. PRICE
A04

ELECTRONIC FORM