To be presented at Intelligent Vehicles '93 Conference, Tokyo, Japan, July 14-16, 1993.

VISUAL SERVOING USING RELEVANT 2-D IMAGE CUES

Daniel Raviv* ** and Martin Herman**

*Robotics Center and Electrical Engineering Department
Florida Atlantic University, Boca Raton, FL 33431

**National Institute of Standards and Technology
Bldg 220, Room B124, Gaithersburg, MD 20899

ABSTRACT

Visual servoing is a mode of navigation in which the motion of a robot vehicle is controlled through servoing on visual cues. Many visual servoing behaviors can be achieved using direct 2-D image cues, without requiring 3-D information reconstructed from the imagery. Usually, only a few *relevant* image cues need to be extracted for each behavior. We present three case studies to demonstrate these ideas: road following, landing on a surface, and visual looming for obstacle avoidance and object approach. For each of these, we show which 2-D information can be used directly to control the robot's motion.

1. INTRODUCTION

An important mode of visual navigation is visual servoing, in which the motion of an autonomous vehicle is controlled through servoing on visual cues. There seem to be a primitive set of robot motion behaviors that can be achieved through visual servoing. Some of these are obstacle avoidance, boundary following, motion relative to environmental objects, approaching an object without collision, pursuing a moving target, and landing on a surface.

In this paper, we argue that in many cases, visual servoing behaviors can be achieved using direct 2-D image cues (both spatial and temporal). These cues serve as inputs to the motion control algorithms. In such cases, an explicit reconstruction of the 3-D scene may not be required. Usually, only a few *relevant* image cues need to be extracted for each behavior. For example, boundary following (as in road following) can be achieved using isolated image information contained in either the inner or outer boundaries of curved roads. As will be shown below, for inner boundaries (which are convex), a feature that may be sufficient for road following is the tangent point on the road inner edge (i.e., the point on the road inner edge lying on an imaginary line tangent to the road edge and passing through the camera) and its image velocity. Therefore, during the boundary following behavior, all image processing effort may be directed towards reliably finding and tracking the tangent point and extracting its image velocity.

In some cases important features do not exist (and are meaningless) in 3-D and can be extracted only in the 2-D image. An example is the vanishing point, which is meaningful only in the 2-D image. As will be shown below, this feature is useful for autonomous landing of an air vehicle. Another example is the focus of expansion, which indicates the instantaneous heading of the camera.

The advantages of this 2-D approach, as will be elaborated below, are that it is simple, fast, and robust. There is no need for 3-D reconstruction. It results in a tight perception-action loop to directly generate action commands from image cues.

The approach taken here derives from several approaches currently in the literature, including "purposive and active vision" [1] and "animate vision" [2]. The common element between these approaches and our approach is that only visual information that can be used specifically and directly for the desired task or behavior is sought.

We present three case studies to demonstrate these ideas: road following, autonomous landing, and visual looming for obstacle avoidance and object approach.

2. SERVOING THE HEADING VECTOR

The heading vector at any instance of time is the instantaneous translation vector of the robot. Many of the primitive navigation behaviors involving visual servoing can be thought of as servoing the heading vector.

Consider the case of road following. We define a *road* as any continuous, extended, curvilinear feature. The goal of road following is to follow along this feature over an extended period of time. In what we normally think of as road following, a road is defined either by its boundaries or by an extended solid or dashed white line. Here, the goal is not only to follow along these features but also to stay within a constant lateral distance from these features.

Figure 1 shows a point on a vehicle and the left-hand side road edge. The unit vector $\hat{h}$ is the instantaneous heading of the vehicle, O is the instantaneous center of curvature of the vehicle path, and r is the instantaneous radius of curvature of this path. Road following is an activity that involves servoing $\hat{h}$ such that it follows the road edge. It is desired that $\hat{h}$ be servoed such that it is always parallel to the tangent to the local curvature of the road edge, and such that the distance s of a point on the vehicle from the road edge is maintained at a constant value.

In the image domain, the heading vector is represented by the instantaneous Focus Of Expansion (FOE), i.e., the 3-D vector extending from the camera focal point to the image FOE point is in the same direction as the heading vector. Therefore, servoing the heading vector is equivalent to servoing the robot's motion using as feedback the position and motion of image features relative to the FOE. We think of this as servoing the FOE in the image. For example, we will show that road following can be achieved using a feedback loop where the position and motion of the road tangent point relative to the FOE are used as feedback variables.

3. MOBILITY CONTROL USING 2-D IMAGE DOMAIN INFORMATION

Many current vision-based mobility control systems use 3-D visually-derived information when making motion control decisions [15-18]. This 3-D information might take the form of range points or range images, 3-D line or surface descriptions, 3-D object descriptions, etc. For example, many existing road following algorithms convert the information extracted from images into a 3-D, vehicle-centered coordinate system, often aligned with the ground plane. Steering, acceleration, and braking decisions are then determined in this coordinate system. A 3-D reconstruction is therefore performed before control decisions are made. In this paper, we show that control algorithms can be developed which *directly* use observable image information represented explicitly or implicitly in the 2-D image sequence. For example, the position and motion of features in the image can be used to determine steering decisions. A 3-D reconstruction is not required.

There are several advantages to this 2-D-based approach.

1. It is more robust. Fewer hypotheses, sensor measurements, and calibrations need to be performed for control when using 2-D information rather than 3-D informa-

tion. The reason is that the problem of converting 2-D image information to 3-D world information requires either hypotheses about the world to be made (e.g., smoothness hypotheses, lighting hypotheses, object hypotheses), calibrations to be made (e.g., stereo calibrations, inertial navigation system calibrations), or many sensor measurements to be made (e.g., inertial navigation system measurements). In using visual information as part of control algorithms, the closer the visual information is to raw data form, the closer it reflects the real world, and the more robust will be the control algorithms based on this information. An example of this is converting image velocity of points to range. Such a process requires that the velocity of the camera be known. However, the measurement of velocity will have errors, which will lead to range errors. Therefore, we expect that it would be more robust not to use velocity to perform navigation control. Further, a control algorithm that does not require velocity is simpler and less expensive.

2. The approach is task-dependent, therefore only relevant information has to be extracted from the images. For example, the road following behavior may only require information dealing with the road tangent point [20]. Other portions of the road edge, although important for extracting the tangent point, may not be necessary as inputs to the steering control algorithm. Each navigation behavior will, in general, need different *relevant* information. Such relevant information may be a set of visible or invisible, spatial or temporal, features. Using only a small set of information from a sequence of images implies simpler and less expensive processes.

3. The closer the relevant visual information is to raw form, the less computation is required to extract this information. Such an approach is therefore simpler and much faster.

Next we present the three case studies.

4. CASE STUDY 1: ROAD FOLLOWING

4.1. COORDINATE SYSTEM

The equations for this case study are defined in a coordinate system which is fixed with respect to the camera on board the vehicle. This coordinate system is shown in Figure 2. We assume that the camera is mounted on a vehicle (later we explain how) moving in a stationary environment. Assume a pinhole camera model and that the pinhole point of the camera is at the origin of the coordinate system. This coordinate system is used to measure angles to points in space and to measure optical flow at these points. We use spherical coordinates (R–θ–ϕ) for this purpose. In this system, angular velocities ($\dot\theta$ and $\dot\phi$) of any point in space, say P, are identical to the optical flow values at P' in the image domain.

4.2. TWO-WHEELED VEHICLE

For our analysis, we use a theoretical two-wheeled vehicle as illustrated in Figure 3. A rigid frame of length $2m$ holds both wheels. A steering wheel angle is applied to both wheels simultaneously, i.e., if one wheel is rotated by an angle β relative to the frame, the other wheel will rotate by the same angle. The camera is mounted such that its pinhole point is located above the front wheel center, and it rotates with the front wheel. The optical axis of the camera coincides with the instantaneous translation vector (heading) of the front wheel.

The following geometrical relationship holds for the vehicle in Figure 3:

$$r = \frac{m}{\sin\beta}. \tag{1}$$

The frame length m is usually known. Thus the instantaneous radius of curvature r of the vehicle path can be determined by measuring the steering angle β.

Figure 4 is an overall description of the system including the spherical coordinate system. For convenience we chose to have the Z axis pointing down. However the same coordinate system as described in Figure 2 is used here. The camera is mounted at some height above the ground and rotates with the front wheel. The position of any point on the road can be expressed with the coordinates R,θ and ϕ, as shown in Figure 4.

In the following analysis, we assume a moving vehicle in a stationary environment. The road is assumed to be planar, and road edges are assumed to be extractable. Figure 5 shows examples of road images obtained from a camera mounted on a vehicle.

4.3. EQUATIONS OF MOTION AND OPTICAL FLOW

We have recently developed a new visual field theory that relates six-degree-of-freedom camera motion to optical flow for a stationary environment [21,22]. This theory provides us with a theoretical and scientific basis for developing constraints, control schemes, and optical flow-based visual cues for road following. This section reviews this theory as it relates to the road following problem.

First we describe 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. Some of the equations can be found in many books, e.g., see [23].

Let the instantaneous coordinates of a 3-D point P (Figure 2) be $\mathbf{R} = (X,Y,Z)^T$ (where the superscript T denotes transpose), the instantaneous translational velocity of the camera be $\mathbf{t} = (U,V,W)^T$, and the instantaneous angular velocity of the camera be $\boldsymbol{\omega} = (A,B,C)^T$. Now consider a specific motion in the instantaneous XY ($\phi = 0$) plane of the camera coordinate system defined by:

$$\mathbf{t} = (U,V,0)^T. \tag{2}$$

$$\boldsymbol{\omega} = (0,0,C)^T. \tag{3}$$

This means that the translation vector may lie anywhere in the instantaneous XY plane while the rotation is about the Z-axis. It can be shown [21] that

$$\begin{bmatrix}\dot\theta\\ \dot\phi\end{bmatrix} = \begin{bmatrix} \frac{-Y}{X^2+Y^2} & \frac{X}{X^2+Y^2} & 0 \\ \frac{-XZ}{\sqrt{X^2+Y^2}(X^2+Y^2+Z^2)} & \frac{-YZ}{\sqrt{X^2+Y^2}(X^2+Y^2+Z^2)} & \frac{\sqrt{X^2+Y^2}}{X^2+Y^2+Z^2} \end{bmatrix} \begin{bmatrix} -U+CY \\ -V-CX \\ 0 \end{bmatrix} \tag{4}$$

As mentioned earlier, $\dot\theta$ and $\dot\phi$ of a point in space (i.e., the angular velocities in the camera coordinate system) are the *same* as the optical flow components $\dot\theta$ and $\dot\phi$.

Consider the case where the optical flow value of $\dot\theta$ is constant. From equation set (4), 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

$$\left[X + \frac{V}{2(C+\dot\theta)}\right]^2 + \left[Y - \frac{U}{2(C+\dot\theta)}\right]^2 = \left[\frac{V}{2(C+\dot\theta)}\right]^2 + \left[\frac{U}{2(C+\dot\theta)}\right]^2. \tag{5}$$

as displayed in Figure 6.

All points in 3-D space that lie on the cylinder described by Equation (5) and which are visible (i.e., unoccluded and in the field of view of the camera) produce the same instantaneous horizontal optical flow $\dot\theta$. We call the cylinder on which equal flow points lie the *equal flow cylinder*.

4.4. ZERO FLOW CYLINDERS

One of the equal flow cylinders corresponds to points in 3-D space that produce zero horizontal flow. We call this cylinder a *zero flow cylinder*. The equation that describes the zero flow cylinder can be obtained by setting $\dot\theta = 0$ in Equation (5), i.e.,

$$\left(X+\frac{V}{2C}\right)^2+\left(Y-\frac{U}{2C}\right)^2=\left(\frac{V}{2C}\right)^2+\left(\frac{U}{2C}\right)^2. \quad (6)$$

We have shown [20] that if the Z component of the camera rotation vector ω is positive (i.e., $C > 0$), then visible points in the XY plane that are inside the zero flow cylinder produce positive horizontal optical flow ($\dot{\theta} > 0$), while visible points outside the zero flow cylinder produce negative horizontal optical flow ($\dot{\theta} < 0$) in the image. If ω is negative (i.e., $C > 0$) then the opposite is true.

4.5. EQUAL FLOW CYLINDERS AS A FUNCTION OF TIME

As the camera moves through 3-D space, the equal flow cylinders move with it. Figure 7 shows sections of equal flow cylinders as a function of time. At each instant of time, the radii of the equal flow cylinders are a function of the instantaneous motion parameters t and ω. The locations of the equal flow cylinders 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 t, and their symmetry axes are parallel to the instantaneous rotation vector ω. (In Figure 7, the direction of ω varies over time.) Each zero flow cylinder lies to the left or right of the translation vector depending on whether the instantanous rotation is positive or negative, respectively.

4.6. ROAD FOLLOWING BY GAZING AT INNER ROAD EDGE

In this section, we consider following along a circular road by gazing at the inner road edge. Given visual cues, a goal of the control system is to find the steering angle. If the vehicle is already on a path that follows the road, then only *changes* in steering angle are necessary. Figure 8 shows a vehicle moving around a circular road of radius l. The path traversed by the vehicle is a circle of radius r. Let the unit vector $\hat{t}$ indicate the direction of the *tangent line*, a line that contains the camera pinhole point and is tangent to the inner road edge. It can be proved [20] that the tangent point T lies on the instantaneous zero flow cylinder if the camera orientation is fixed relative to the vehicle. We can also prove that no matter how far the vehicle is from the inner road edge, the tangent point lies on the zero flow cylinder. Thus the horizontal component of optical flow of the tangent point is always zero.

Note that although the tangent point appears as a feature point in the image, it is not generated by a single point in space. Rather, the points in space that generate this feature constantly change. We call this point a *virtual feature point*.

In Figure 8, therefore, the optical flow $\dot{\theta}$ due to point T is zero. Let the distance from the vehicle to the inner road edge be s, and let θ be the positive angle to $\hat{t}$ measured from the X-axis. From Figure 8, the following relationships hold:

$$l = r\sin\theta, \quad (7)$$

$$s = r - l = r(1-\sin\theta). \quad (8)$$

Differentiating Equation (7) with respect to time:

$$\dot{l} = \dot{r}\sin\theta + r\dot{\theta}\cos\theta \quad (9)$$

where dot denotes derivative with respect to time. For a circular road, l is constant, and thus $\dot{l}$ can be set to zero in Equation (9):

$$0 = \dot{r}\sin\theta + r\dot{\theta}\cos\theta$$

$$\dot{r} = -r\,\dot{\theta}\cot\theta. \quad (10)$$

When the vehicle is moving on a perfect circular path both $\dot{r}$ and $\dot{\theta}$ are equal to zero. However, suppose the vehicle's path is not a perfect circle. Since r is the instantaneous radius of curvature of the vehicle motion, $\dot{r}$ is the rate at which the curvature changes. Equation (10) suggests a way of controlling the vehicle motion so as to achieve a constant circular motion. Consider the two-wheeled vehicle described above. From Equation (1), we can derive the following:

$$\beta = \sin^{-1}\left(\frac{m}{r}\right). \quad (11)$$

Equation (11) gives a value of the steering angle β as a function of the instantaneous radius of curvature r and the distance $2m$ between the two wheels. Normally the value m is known. For a more realistic vehicle (such as a four-wheeled vehicle with front-wheel steering), some other relationship will hold.

In Equation (10), $\dot{r}$ is the rate at which the radius of curvature of the vehicle motion is changing. We can express $\dot{r}$ as a function of the steering angle β by differentiating Equation (1) with respect to time:

$$\dot{r} = \frac{-m\cos\beta}{\sin^2\beta}\,\dot{\beta}. \quad (12)$$

Substituting Equations (12) and (1) into (10) and solving for $\dot{\beta}$:

$$\dot{\beta} = \dot{\theta}\tan\beta\cot\theta. \quad (13)$$

Equation (13) suggests a partial control scheme whose inputs are the current steering angle β, the current angle θ of the tangent line relative to the X-axis, and the optical flow $\dot{\theta}$ of the tangent point. All of these inputs can be measured. The variable being computed is the rate of change of the steering angle, $\dot{\beta}$. Equation (13) provides the gain $\tan\beta\cot\theta$ by which $\dot{\theta}$ should be multiplied in order to get the correct change in steering wheel angle. This gain depends on the current steering wheel angle β and the angular location θ of the tangent point in the image. Note that the change in steering control command should be the negative of the value of $\dot{\beta}$ derived in Equation (13).

Figure 5 shows a sequence of images taken from a camera mounted on a vehicle. The images in the figure are numbered in the same order in which they were taken. The road is almost circular. Note that the tangent point stays (almost) at the same location in each image in the sequence. If the road were perfectly circular and the vehicle were moving on a perfect circular path, then the position of the tangent point would not change from image to image. However, if the vehicle's path were not a perfect circle, then its steering can be controlled by measuring horizontal changes in the position of the tangent point. These changes are the horizontal component of optical flow at that point, and can be used to generate changes ($\dot{\beta}$) in the steering wheel command β.

Note that Equations (1) and (11) hold only for certain types of vehicles. Vehicles with other wheel and steering configurations will result in different expressions relating steering angle to the radius of curvature of motion. In all such expressions, however, there should be a one-to-one relationship between β and r. These expressions can then be substituted into Equation (10) to derive the relevant control signals. It is important to emphasize that the derivation of $\dot{\beta}$ takes into account the kinematics of the system but *not* the dynamics. This is also the reason why we emphasize that the control scheme is not complete.

If the rate of change of the steering angle, $\dot{\beta}$, is the only variable being controlled (as indicated in Equation (13)), then in practice the vehicle may not maintain a constant distance from the edge of the road. Therefore, in addition to Equation (13), Equation (8) can also be used to control the vehicle to achieve a constant circular motion. Substituting Equation (1) into (8) and solving for β:

$$\beta = \sin^{-1}\left[\frac{m}{s}(1-\sin\theta)\right]. \quad (14)$$

Equation (14) suggests a partial control scheme whose inputs are the measured angle θ of the tangent line relative to the X-axis, the desired distance s of the vehicle from the inner road

edge, and the distance $2m$ between the front and rear wheels. The variable being computed is the steering angle β.

The control signals (β and $\dot{\beta}$) and partial control schemes suggested above assume that the road is circular, that the center of curvature of the desired vehicle path coincides with the center of curvature of the circular road, and that the road is planar. It is also assumed that the tangent point (in the image) can always be determined, and that the vehicle heading coincides with the camera optical axis. The main significance of this approach is that (1) it shows that the tangent point and its optical flow are important visual control signals for road following, and that potentially they are sufficient to achieve road following, (2) a tight perception-action loop is possible for road following which is simple and therefore computationally inexpensive, (3) image information can potentially be used *directly* as input into a control algorithm, (4) only a few measurements may be needed to control the vehicle, (5) the approach is independent of the speed of the vehicle, (6) it is independent of the camera height above the road, and (7) only a very small portion of the image -- the portion around the tangent point -- may need to be analyzed, in principle. (Of course, item (7) may not be true in practice since larger portions of the road may have to be extracted in order to reliably find the tangent point.)

Suppose we want to follow a road which is curved but not circular. Figure 9 shows two cases. In Figure 9a the radius of curvature increases as the vehicle moves. In Figure 9b, the radius of curvature decreases. In Figure 9a, the tangent point lies on some equal flow cylinder whose $\dot{\theta}$ optical flow is negative. In Figure 9b, the tangent point lies inside the zero flow cylinder, and its $\dot{\theta}$ optical flow is positive. Therefore, intuitively, if the horizontal component of the optical flow, $\dot{\theta}$, at the tangent point is measured, then its value can be used as a control signal for steering the vehicle. If $\dot{\theta}$ is negative (Figure 9a) then the steering command is to increase the radius of curvature of the vehicle's current motion. If $\dot{\theta}$ is positive (Figure 9b), then the steering command is to decrease the radius of curvature of the vehicle's current motion by sharpening the turn.

4.7. ROAD FOLLOWING BY GAZING AT OUTER ROAD EDGE

We have has thus far considered road following using information about the inner edge of curved roads. In this section, we look at the problem of road following using information about the outer edge of curved roads. Figure 10 shows the outer edge of a circular road. In the figure, a vehicle is following a road whose outer edge has radius l_o. The path traversed by the vehicle is a circle of radius r. Consider the direction $\theta = \alpha$ in the camera. Refer to the definition of θ in Figure 4. This angle determines a direction defined by the plane $\theta = \alpha$. The intersection of this plane with the outer road edge is at point V. The ray from the camera to point V is (α, ϕ) for some ϕ. Figure 10 shows the top view, where a is the horizontal projection of the line connecting the camera pinhole point with point V.

Let h_c be the height of the camera above the ground. During driving, this height is constant. Therefore

$$a = \frac{h_c}{\tan\phi}. \tag{15}$$

In this section, we show that road following can be achieved, in principle, from information at a single point on the outer road edge, the point V. During driving, for constant azimuth direction α, the elevation angle ϕ can change as the value a changes. The point V is a virtual feature point and the value α, which defines this point for a given road, is arbitrarily chosen.

Let the distance from the vehicle to the outer road edge be s_o. From Figure 10, the following relationships hold:

$$l_o^2 = r^2 + a^2 - 2ra\cos\alpha, \tag{16}$$

$$s_o = l_o - r. \tag{17}$$

Combining Equations (15) and (16) yields:

$$l_o^2 = r^2 + \frac{h_c^2}{\tan^2\phi} - \frac{2rh_c\cos\alpha}{\tan\phi}. \tag{18}$$

Combining Equations (17) and (18) and solving for r:

$$r = \frac{h_c^2 - s_o^2\tan^2\phi}{2\tan\phi\,(s_o\tan\phi + h_c\cos\alpha)}. \tag{19}$$

Substituting Equation (1) into (19) and solving for β:

$$\beta = \sin^{-1}\left[\frac{2m\tan\phi\,(s_o\tan\phi + h_c\cos\alpha)}{h_c^2 - s_o^2\tan^2\phi}\right]. \tag{20}$$

This equation suggests a partial control scheme whose inputs are the arbitrarily chosen azimuth angle α, the measured elevation angle ϕ of the virtual feature point V, the desired distance s_o of the vehicle from the outer road edge, the known distance $2m$ between the front and rear wheels of the vehicle, and the known height h_c of the camera. The variable being computed is the steering angle β.

Similarly, equations of the following kind can be found for $\dot{r}$ and $\dot{\beta}$,

$$\dot{r} = f\,(\dot{\phi}, \phi, r, \alpha, h_c), \tag{21}$$

$$\dot{\beta} = g\,(\dot{\phi}, \phi, \beta, \alpha, h_c, m). \tag{22}$$

Equation (22) is for the two-wheeled vehicle described above, and may be formed by substituting Equations (1) and (12) into (21) and solving for $\dot{\beta}$.

Equations (20) and (22) can be used to achieve constant circular motion. However, suppose we want to follow a road which is curved but not circular. In this case, the radius of curvature may either increase (Figure 9a) or decrease (Figure 9b) as the vehicle moves. We can use the rate $\dot{\phi}$ at which the position of the outer road edge changes as measured in the $\theta = \alpha$ plane. For increasing radius of curvature, $\dot{\phi}$, will be negative, while it will be positive for decreasing radius of curvature. This value can therefore be used as a control signal for steering the vehicle. If $\dot{\phi}$ is negative, then the steering command is to increase the radius of curvature of the vehicle's current motion. If $\dot{\phi}$ is positive, then the steering command is to decrease the radius of curvature of the vehicle's current motion by sharpening the turn.

5. CASE STUDY 2: AUTONOMOUS LANDING

5.1. THE VANISHING POINT

This section discusses how the vanishing point of a runway may be used for the autonomous landing task. The vanishing point is a visual cue which is relevant primarily for highly structured environments; the runway is an example. This cue does *not* exist (and is meaningless) in 3-D and can be extracted in 2-D only. In fact, this cue does not even exist in the 2-D image; it must be extracted from other visible 2-D features. By using an active vision approach (in particular a fixating camera) and this visual cue, some vision-based control computations can be significantly simplified. The idea behind the active vision landing approach is based on fixating the camera at the vanishing point of the projection of the runway onto the image.

The vanishing point of the imaged runway is the point in the image plane where the projection of the two parallel side lines of the runway intersect. Obviously, this intersection point does not exist in a finite 3-D world and cannot be "reconstructed." Since the vanishing point is a projection of a point "located" at infinity, any translational motion of the camera

does *not* change the location of the vanishing point in the image plane. However, rotational motions of the camera will, in general, cause the vanishing point to shift in the image plane.

5.2. VANISHING POINT AND FIXATION

The runway vanishing point provides information about aircraft orientation which can serve as a key visual feature for autonomous landing. If the camera is continuously fixated at the vanishing point through a visual feedback loop such that the optical axis of the camera contains the vanishing point, then there will be no pitch and yaw rotational motions of the camera The camera may, however, rotate about its roll axis. This also means that if the vanishing point does shift in the image, it is due to pitch and/or yaw rotational motions only.

When the camera undergoes translation along its optical axis, the vanishing point becomes the Focus of Expansion (FOE). For arbitrary camera translations along the X, Y and Z axes of the camera the vanishing point is always *stationary* in the image.

Next we shall explain how the vanishing point, as it evolves as a function of time, can be used to extract valuable information about aircraft orientation, which can be used for landing. Suppose the camera is mounted on the aircraft and it has independent pan/tilt control. Then by maintaining fixation on the vanishing point it is possible to:

1. Measure the two-degree-of-freedom orientation (pitch and yaw) of the aircraft relative to the fixating camera by measuring the camera pan/tilt angles, thus allowing the measurement of the aircraft 2-D orientation relative to the runway. (Note that when the camera fixates at the vanishing point the optical axis becomes parallel to the longitudinal axis of the runway.)
2. Reduce the need for other rotation sensors (e.g., gyroscopes).

Practically, during the fixation process, the vanishing point which is being used to close the fixation loop may shift from the optical axis due to computation delay and loop response. In this case, it may be useful to add the velocity of the vanishing point, i.e., its optical flow, (in addition to its location) relative to the optical axis to generate control signals to maintain fixation (and the resulting camera stabilization).

Other landing related visual cues that are based on fixating the camera on the runway vanishing point have also been developed and are discussed in [24].

5.3. CONTROL LOOP BASED ON THE VANISHING POINT

Figure 11 shows possible control loops for maintaining the vanishing point at the center of the image. The desired location of the vanishing point is the input to the closed loop system. The feedback signals are the location of the vanishing point in the image and its velocity (flow). The parameters being controlled are the pan and tilt of the camera. These control loops allow us to use the fixating camera system as a gyroscope, thus aiding in the autonomous landing task.

6. CASE STUDY 3: VISUAL LOOMING

6.1. THE VISUAL LOOMING EFFECT

The visual looming effect deals with the expansion of an object's size in the image. Usually, this expansion is a result of a decrease in the relative distance between the camera and the viewed object. Visual looming is a quantitative temporal feature which can be very useful for generating control signals for obstacle avoidance or for approaching an object without collision.

The perception of looming is critical to survival for creatures of nature, since it is an indication for possible collision. 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.

Although looming and looming-related actions have been studied by many researchers (primarily psychologists), most of the work is qualitative or limited-quantitative. In order to use the existing results for robotics applications, a more mathematically-oriented approach for looming is needed. The following sections present such an approach.

6.2. QUANTITATIVE LOOMING

Assume an observer is moving in a 3-D environment filled with several balls (Figure 12). During the motion of the observer relative to the balls, both the size of each ball and the location of 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.

Lee [4] 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. Our definition of looming is related to, but different from, the time-to-contact. The time-to-contact according to [4] 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)*.

We now introduce a mathematical definition of looming, and a way for measuring looming, using expansion of the projection of objects on the retina [14].

6.3. MATHEMATICAL DEFINITION OF LOOMING

Let a camera and an object (e.g., a ball) move arbitrarily in a 3-D environment. 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. 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. 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.,

$$L = -\lim_{\Delta t \to 0} \frac{\frac{(R_2-R_1)}{\Delta t}}{R_1}, \qquad (23)$$

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

$$L = -\frac{\dot{R}}{R}. \tag{24}$$

where dot denotes derivative with respect to time. The reason for the *negative sign* in Equations (23) and (24) 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 of a point.

6.4. DIRECT CALCULATIONS OF LOOMING

This approach deals with relative motion between the camera and the object. It is based on measuring the looming at the image-region (or object) level, rather than at the pixel level. Techniques for measuring looming at the pixel level are provided in [14].

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 13 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 $|\dot{t}|$, the radius of the ball r, and the distance from the center of the ball to the observer R. At $t=t_1$

$$\sin\beta = \frac{r}{R}. \tag{25}$$

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

$$\sin(\beta+\Delta\beta) = \frac{r}{R-|\dot{t}|\Delta t}. \tag{26}$$

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

$$\sin(\beta+\Delta\beta) = \frac{r}{R}\,(1+\frac{|\dot{t}|}{R}\Delta t). \tag{27}$$

By subtracting Equation (25) from Equation (27), then dividing by Δt, and letting $\Delta t \to 0$, we obtain:

$$\frac{d\sin\beta}{dt} = \frac{r|\dot{t}|}{R^2}. \tag{28}$$

Using the relations $\frac{r}{R} = \sin\beta$, $\frac{|\dot{t}|}{R} = -\frac{\frac{dR}{dt}}{R}$ and $\frac{d\sin\beta}{dt} = \cos\beta\frac{d\beta}{dt}$ in Equation (28) we obtain:

$$\frac{\frac{dR}{dt}}{R} = -\frac{\frac{d\beta}{dt}}{\tan\beta}. \tag{29}$$

and the expression for looming becomes:

$$L = -\frac{\dot{R}}{R} = \frac{\dot{\beta}}{\tan\beta} = \frac{d}{dt}\ln(\sin\beta), \tag{30}$$

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 (see [14]) that a retina which is constructed in a logarithmic fashion, i.e., where 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 the ball is simple. Relative change of area is also proportional to looming. We shall show next that

$$\frac{\frac{dA}{dt}}{2A} = -\frac{\frac{dR}{dt}}{R} = L, \tag{31}$$

where A is the projected area of the ball on the 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 13, 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 (30)) to be the expression for looming. Thus the relative change of the projected area can be used to approximate looming values using Equation (31).

7. CONCLUSIONS

In this paper, we have demonstrated how some visual servoing behaviors can be achieved using direct 2-D image cues (both spatial and temporal). These cues serve as inputs to the motion control algorithms. In such cases, an explicit reconstruction of the scene is not required. We have also demonstrated that each particular behavior will, in general, need different relevant information from the imagery.

The advantages of this 2-D based approach are that it is simple, fast and robust. There is no need for 3-D reconstruction. It results in a tight perception-action loop to directly generate action commands from image cues.

8. ACKNOWLEDGEMENT

This work was supported in part by a grant to Florida Atlantic University from the National Science Foundation, Division of Information, Robotics and Intelligent Systems, Grant # IRI-9115939.

REFERENCES

[1] Aloimonos, J. "Purposive and Qualitative Active Vision." *Proc. DARPA Image Understanding Workshop*, September 1990.

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

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

[14] Raviv, D., "A Quantitative Approach to Looming", NIST Internal Report, NIST-IR-4808, 1992.

[15] Dickmans, E. and Grafe, V. "Applications of Dynamic Monocular Machine Vision." *Machine Vision and Applications*, Vol. 1, 1988.

[16] Dickmans, E. and Grafe, V. "Dynamic Monocular Machine Vision." *Machine Vision and Applications*, Vol. 1, 1988.

[17] Thorpe, C., Hebert, M., Kanade, T. and Shafer, S. "Vision and Navigation for the Carnegie-Mellon Navlab." *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 10(3), 1988.

[18] Turk, M., Morgenthaler, D., Gremban, K. and Marra, M. "VITS – A Vision System for Autonomous Land Vehicle Navigation." *IEEE Trans. on Pattern Analysis and Machine Intelligence*, May 1988.

[20] Raviv, D. and Herman, M., "A New Approach to Vision and Control for Road Following", *IEEE Workshop on Visual Motion*, Princeton, NJ, October 1991.

[21] Raviv, D. "A Quantitative Approach to Camera Fixation." *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Maui,

Hawaii, June 1991.

[22] Raviv, D. and Herman, M. "Towards an Understanding of Camera Fixation." *Proc. 1990 IEEE International Conference on Robotics and Automation*, Cincinnati, Ohio, May 1990, 28-33.

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

[24] Yakali, H.H., and Raviv, D., "A Vision-Based Method for Autonomous Landing", *SPIE Conference on Intelligent Robots and Computer Vision*, Boston, MA, November 1991.

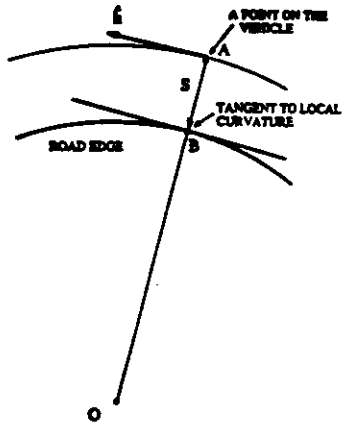

Figure 1: Road Following

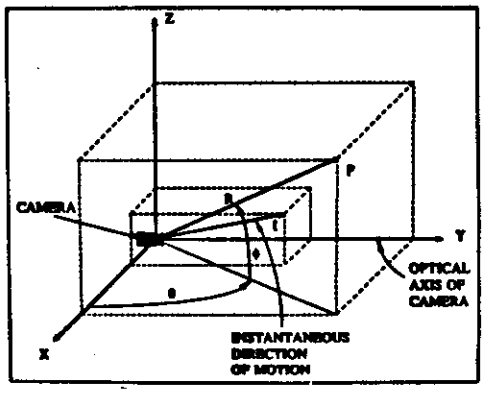

Figure 2: Coordinate System Fixed To The Camera

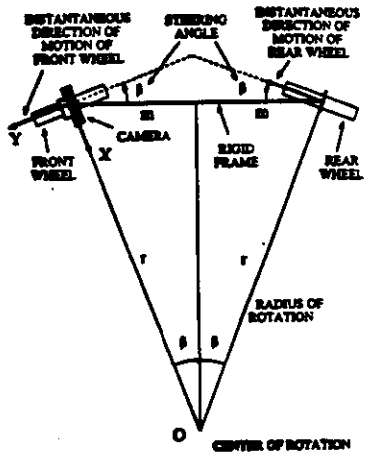

Figure 3: Two-wheeled Vehicle With A Camera

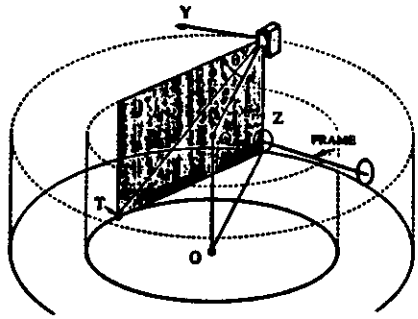

Figure 4: Overall Description Of The System

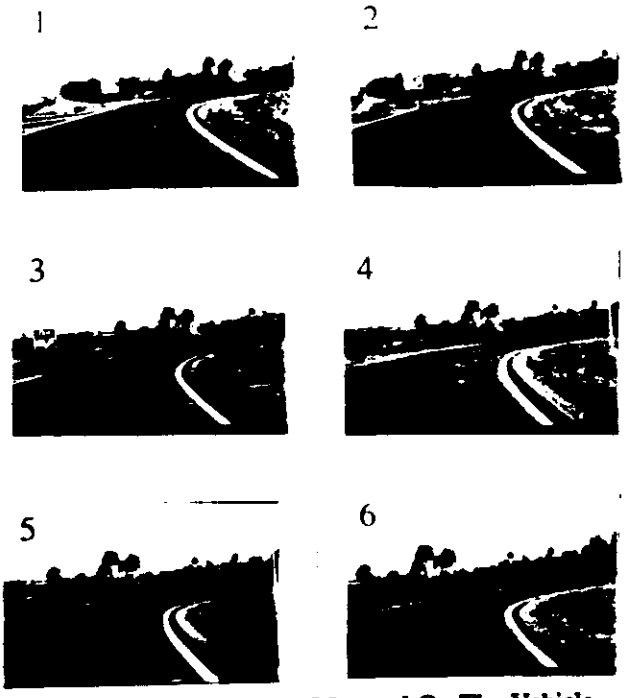

Figure 5: Images Obtained From A Camera Mounted On The Vehicle

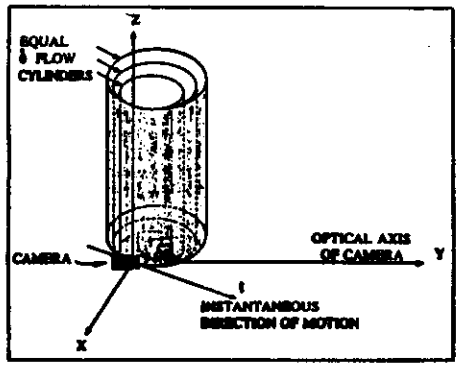

Figure 6: Equal $\dot{\theta}$ Flow cylinders

Figure 7: Sections Of Zero flow Cylinders As A function Of Time

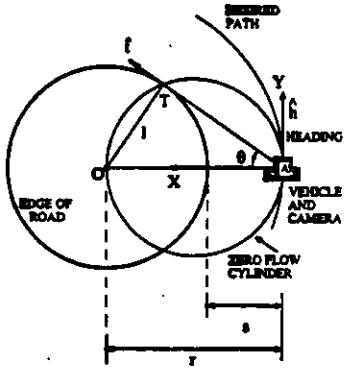

Figure 8: Circular Edge: Top View

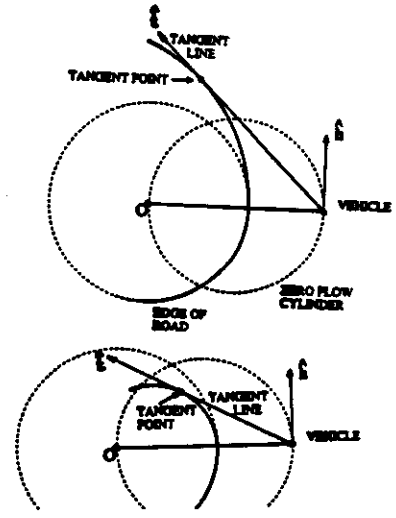

Figure 9: Road Following:
(a) Increasing Radius Of Curvature
(b) Decreasing Radius Of curvature

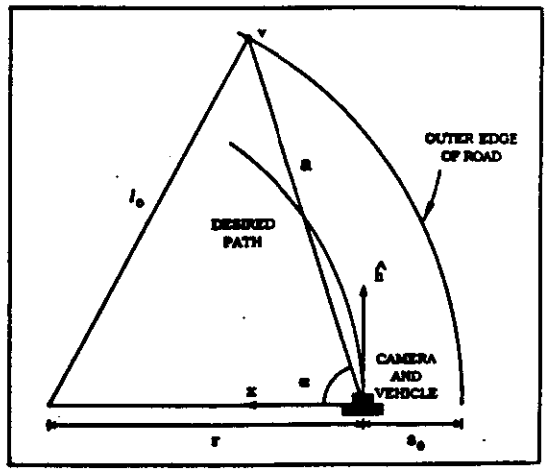

Figure 10: Outer Edge Road Following

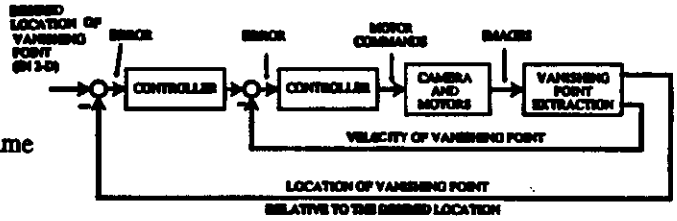

Figure 11: Control Loops For Fixating the Camera At The Vanishing Point

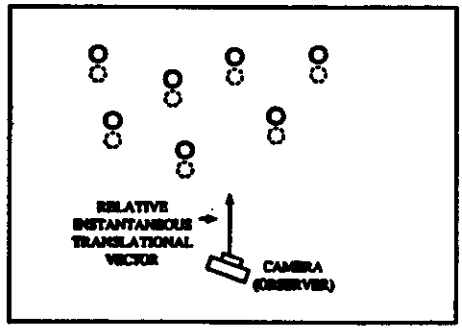

Figure 12: Relative Motion Between A Camera And 3-D Objects

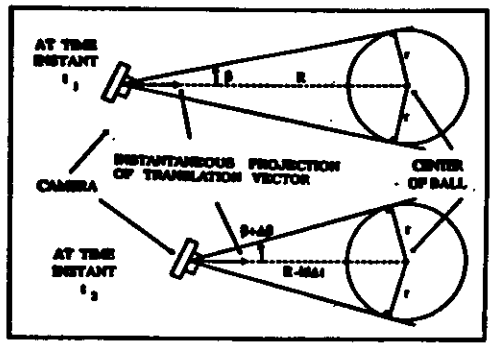

Figure 13: Looming Computations