USING AN ACTIVE VISION SYSTEM TO COMPUTE TIME-UNTIL-IMPACT

Jorge Dias, Jorge Batista, Helder Araújo, Anibal T. de Almeida
Institute of Systems and Robotics
Department of Electrical Engineering
University of Coimbra
Largo Marquês de Pombal, 3000 Coimbra, Portugal
fceljorge@uc.pt

Abstract -- This paper describes the use of an experimental system for active vision to compute time until impact. The time-until-impact is evaluated by using the principle of the human visual fixation process. The mechanical system enables the positioning of the cameras´ support platform by means of three rotational degrees of freedom and also the control of the vergence angle of each camera. The motors used are stepper motors which are controlled by a system specifically designed and built for this purpose. Each camera´s lens has three degrees of freedom, corresponding to zoom, focus and aperture. All the control system electronics as well as the system status data acquisition electronics are integrated in a single control unit. This control unit is connected to a system responsible for the active vision system motion generation.

INTRODUCTION

One of the goals of an active vision system is to enable the processing of scene regions deemed important for the global behaviour of the system. Active vision has implied a deep change on the principles and models used by Computer Vision. That is, a vision system isn' t considered anymore as an isolated system, but instead as a system that cooperates with the mechanical system that controls several degrees of freedom: position in space, vergence, zoom, focus, aperture, and other. One example of such a behaviour is the control of the focus of attention of the visual system. To implement such a type of behaviour, on a dynamic environment, one needs to control the different degrees of freedom so that the region of interest remains, for example, in the central region of the image. This type of control is significantly easier if the feedback process uses the image of the region around the point of intersection of the two cameras' axis. Another example of a behaviour is the adaptation of the system to changes of the scene's illumination. In this case the lens´ aperture has to be accordingly changed. Focussing and zooming are also required because it is also necessary that the objects and other elements in the scene remain within focus.

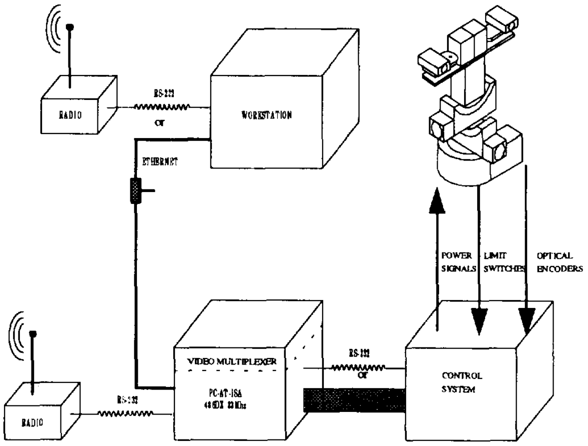

Figure 1- Blocks diagram of the active vision structure

Ballard [4][5], [2], Bajcsy [3], Abbot [1] and others have proved several advantages on the use of Active Vision, when compared with approaches based on the use of fixed camera systems. This concept is linked to the notion of active observer that undertakes a specific type of behaviour to obtain a scene perception well-suited for his purpose. In terms of computer vision this behaviour partially corresponds to the manipulation of the geometric constraints that model perception.

In this paper we describe an inexpensive system for performing active vision experimentation and its use to compute time-until-impact. The system is integrated on the ROBUTER mobile platform of our laboratory.

MECHANICAL STRUCTURE

The mechanical structure of the system developed at our department (Electrical Engineering Department of the University of Coimbra), measures 295mm x 260 mm x 315

0-7803-0891-3/93$03.00 ©1993IEEE

mm, weighting 8.9 Kg-- see Figure 2.

The cameras´ support platform has three rotational degrees of freedom. Each degree of freedom has two limit switches and an incremental optical encoder. The optical encoder is used as a feedback sensor. There is also a reference position, which is required for calibration. The mechanical structure is made up of two blocks: the cameras´ platform and the block of the rotational degrees of freedom. The cameras´ platform enables the vergence motion of the cameras. The block on which the cameras´ platform is mounted has three degrees of freedom whose rotation axis are mutually orthogonal. One of them has a vertical axis of rotation enabling movements in azimuth, whereas the other two have horizontal rotation axis, enabling movements in elevation. The azimuth degree of freedom has a range of 340^{o} (+/- 170^{o}) whereas the other two have a range of 90^{o} (+/- 45^{o}). Each one of these degrees of freedom has a precision of 0.005o. Stepper motors are used as actuators, and have speeds that can vary between 11.35^{o}/sec and 129^{o}/sec.

Figure 2- Mechanical structure of the active vision system

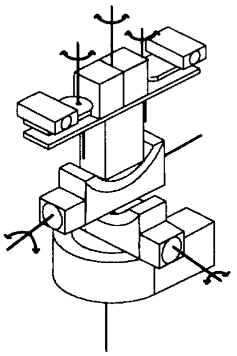

Figure 3 - Degrees of freedom of the mechanical system

Controls of both vergence and intra-ocular distance are performed by means of two stepper motors. Each stepper motor is coupled to each camera´s support by means of gears. This way it is possible to control the vergence angle of each camera with a precision of 0.6^{o} . Each vergence angle has a maximum range of 110^{o} (+/- 55o). This implies that the intra-ocular distance may vary between 10cm and 29cm.

HARDWARE OF THE CONTROL SYSTEM

The hardware is made up of AT-BUS boards. There are two different types of boards: the Control Board (C.B.) and and the Power Supply Board (P.S.B.). Each P.S.B. can supply power to up to three stepper motors, whereas the C.B. can control up to six degrees of freedom. This set of boards was designed to be modular . With these boards it is possible to have a single system controlling 48 stepper motors.

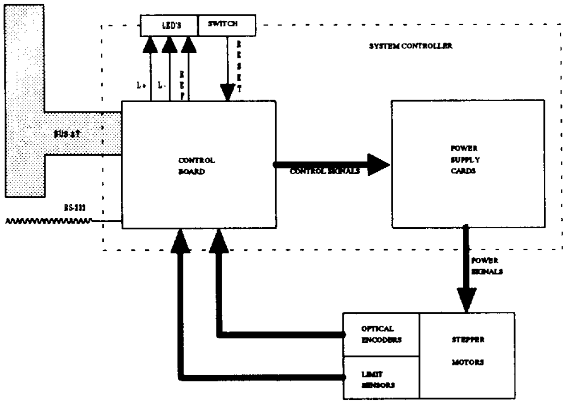

Figure 4 - Controller harware of the active vision system

This control system can communicate either via a RS-232 series channel or via AT-BUS. On the C.B. there is a special purpose processor (CY233 from Cybernetics Microsystems) which is in charge of the serial communications. By using the AT-BUS it is possible to control and monitor, in real time, all the six motors connected to the board. The control of the stepper motors is performed by special purpose processors (CY545 from Cybernetics Microsystems). The C.B. has a local memory of 32 Kbytes, which permits that each CY545 executes a specific program. Besides the information from the limit switches and the optical encoders these boards also use the information provided by direction discriminators (THCT 12024). These direction discriminators are connected to the optical encoders.

The P.S.B. are built around the GS-D200 (from SGS Thomson), which are the chips that power the bipolar stepper motors windings.

The zoom, focus and aperture of each camera are actuated by variable speed DC motors. Feedback on each one of these degrees of freedom is provided by an absolute position potentiometer. Communication with the board that implements the control of these degrees of freedom can also be done either by a series channel or via AT-BUS. The output of each potentiometer is digitized by an 8 bit ADC. The speed of each motor is controlled by changing the frequency of a sequence of pulses of constant width.

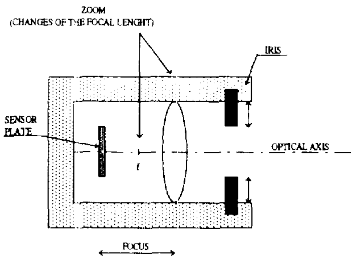

Figure 5 - Degrees of freedom on the motorized lens (zoom, iris and focus)

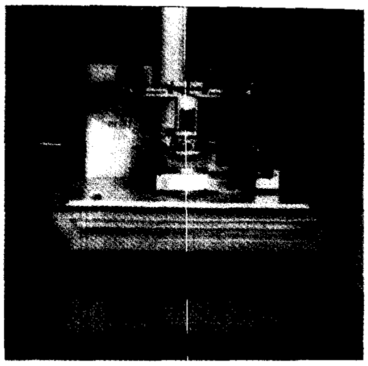

Figure 6 - Active Vision System mounted on the mobile robot

SYSTEM CONTROLLER SOFTWARE

The system controller can communicate both via Ethernet and wireless modem. Currently, the controller software has a set of functions enabling the control of the information flow between the system controller (where basic behaviours are implemented) and the mechanical system controller, and the acquistion of the acceleration and speed parameters of the stepper motors as well as the joint positions. It is also possible to have all the information concerning vergence angles, intra-ocular distance and the position and orientation of the cameras' support platform. Finally it is also possible to have 64X64 images from the left and right cameras. New modules are being added to the system.

COMPUTING TIME-UNTIL-IMPACT

One important aspect of active vision is the advantage of using a frame on the fixation point. Moving a camera a little forward while fixating an environment point in a static scene, objects in front and behind of the fixation point have apparent motions with opposite directions. The apparent velocity is proportional to the distance from the fixation point [5] The results obtained by the experiments made by Ballard show that the computations are simpler than fixed camera vision, as first noted by [2]. Under these conditions we can compute scaled depth (depth per fixation depth).

Using fixation one can compute the time until impact. Consider a simple binocular vision system with two referential associated to the cameras--see figure 7(a). The referentials have origin on the center of the lens, the Z axes have the direction of the optical axis and the Y axes are parallel. Consider a translational motion of the system in which the rotation around X axis is equal a zero (i.e., a motion trajectory that lies on the ZX plane performed by stabilizing on the fixation point). At time T the angle θ represents the rotational angle around an axis orthogonal to the plane of motion, whereas γ represents the direction of the translational axis.

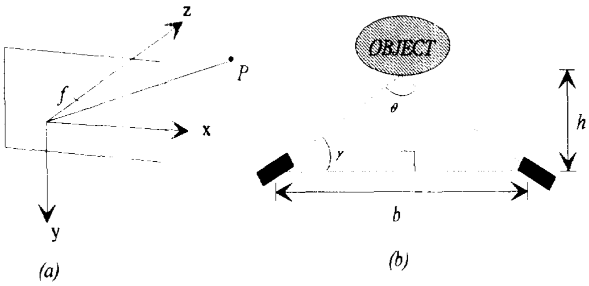

Figure 7 - (a) Schematic representation of the camera model. (b) By measuring the rate of change of the vergence angle we can obtain the time until impact.

By using similar development, the time until impact of the active vision support against the fixated object can be calculated. By measuring the rate of change of the vergence angle we obtain the velocity of the mobile support directly, considering acceleration null - figure 7 (b).

Consider a situation, similar to the figure 7, with an object in front of the active vision system at the distance h. Since the relations between the angles are given by

$$\theta = 180^\circ - 2\gamma \qquad (1)$$

and by using the figure we have

$$h = \frac{b}{2}\tan\gamma \qquad (2)$$

Taking the derivative of the expression above we have

$$\dot{h} = \frac{b}{2}(\sec^2\gamma)\dot{\gamma} \tag{3}$$

and the time until impact is given by

$$tui = \frac{h}{\dot{h}} = \frac{\cos\gamma\sqrt{1-\cos^2\gamma}}{\dot{\gamma}} \tag{4}$$

The time-until-contact computation presupposes the execution of the fixation process by the active vision system, while the object is moving. To perform the fixation process we need to control the system to maintain the object's gravity center in the middle of the image. The object's center of gravity is obtained by preprocessing the images acquired by the system. The preprocessing uses a small window of 64 x 64 pixels from the original image of 512x512 pixels. That window is centered in the image and is converted to a binary image. The processing is fast enough to be included in the fixation control cycle.

To control the system is necessary to maintain the successive positions of the object in the images. For that, the gravity center of the object is continuously tracked by using a Kalman filtering approach. This tracking approach is based on a prediction and matching step and provides a reasonable mechanism to control the matching process and the vergence angle. Kalman filtering is a statistical approach to estimate a time-varying state vector $\mathbf{x}_k$ from noisy measurements $\mathbf{y}_k$ in the instant k. The Kalman filtering is a recursive estimation scheme designed to obtain a prediction of $\mathbf{x}_{k+1|k}$ using the measurements up to the instant k, by matching the dynamic system model, the statistics of the error between the model and reality, and the uncertainty associated with the measurements.

In our real-time application, two Kalman filters are applied by each image and independently to each parameter (u_m, v_m) that describes the center of gravity of the object. Each state vector is of dimension 3 and composed by the position of the given parameter x, its velocity $\dot{x}$ and its acceleration $\ddot{x}$. Therefore, we deal with the following state vectors: $[u_m, \dot{u}_m, \ddot{u}_m]$ and $[v_m, \dot{v}_m, \ddot{v}_m]$. Initially, no information available exists about the kinematics of the image point. For this instant we assume that the velocities and accelerations are zero and the correspondent uncertainties are very large, indicating that we do not trust these guests This is the *bootstrap* stage and it has a duration of three cycles.

The model of the system dynamics is

$$\mathbf{x}_{k+1} = \mathbf{A}\,\mathbf{x}_k + \xi_k \tag{5}$$

with the matrix A given by

$$\mathbf{A} = \begin{bmatrix} 1 & \Delta t & \frac{\Delta t^2}{2} \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix} \tag{6}.$$

The measurement model used is

$$\mathbf{y}_k = \mathbf{C}\,\mathbf{x}_k + \eta_k \tag{7}$$

where $\mathbf{C} = [1 \quad 0 \quad 0]$ and ξ and η are gaussian white noise sequences with covariance matrices known.

The fixation and time-until-impact algorithm can be described as follows:

```
Algorithm
1. Initialize the system;
2. Detect the object;
3. Bootstrap the Kalman filters;
4. Do Forever
        4.1. Predict next position;
        4.2. Control the vergence based
                on the prediction;
        4.4. Detect new object;
        4.5. Match the object and
                actualize the Kalman filters;
        4.6. Estimate the time-until-colision
end
```

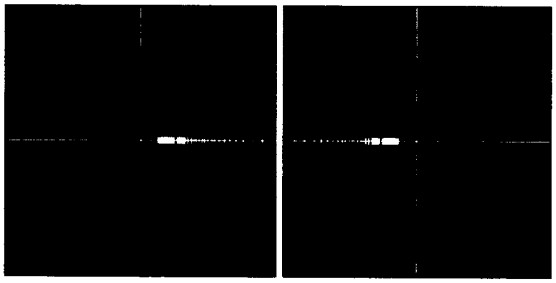

Figure 8- The figure illustrates the right and left images of the fixated object. In this example the object has one trajectory of collision, against to the active vision system. The object is continuously tracked by using Kalman filtering and the images represent the projections of the object position estimation during its movement. By using these estimations it is possible to have one estimation of the time-until-impact of the object.

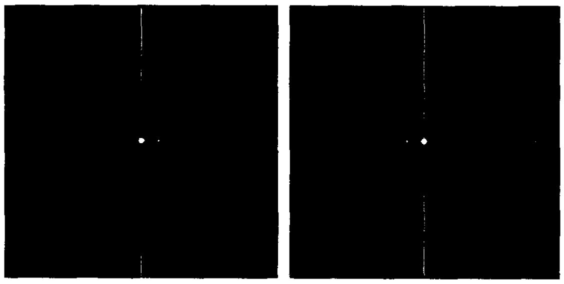

Figure 9 - The figure illustrates the projection of the object in the left and right images during the fixation process.

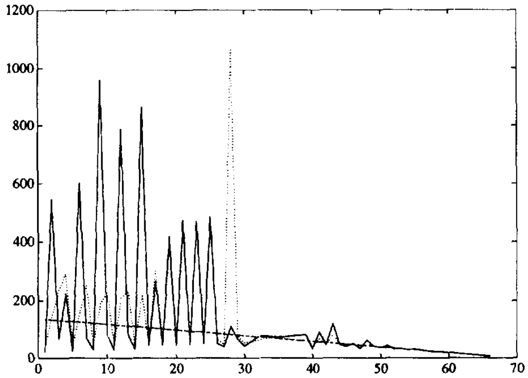

Figure 10 - The figure illustrates the time-until-impact values. The dashed line represents the real values. There are two lines represent the estimated time-until-impact. The first values of these two curves are very unstable because the derivative is approximated by difference between values given by the Kalman filter and the filter is not correctly tracking the point in the first iterations.

CONCLUSIONS

An inexpensive system for active vision experimentation was described. The system has a block with three rotational degrees of freedom, on the top of which a platform for camera support is mounted. The camera support enables the control of the vergence angle of each camera. The actuators of the three rotational degrees of freedom as well as of the vergence angles are stepper motors with incremental optical encoders feedback. The zoom, focus and aperture of each camera can also be controlled. An experiment for time until impact computation by using this experimental system was also described.

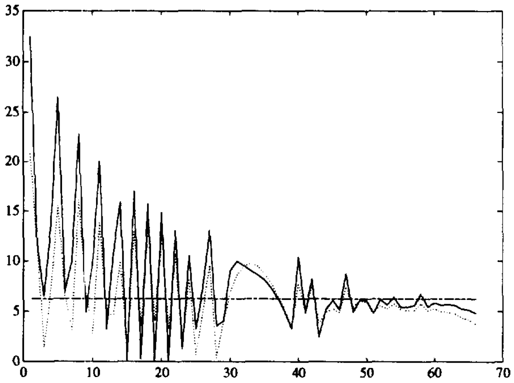

Figure 11 - The diagram shows the curves of the vergence velocity values during the fixation process. In this case the object has constant velocity. The dashed line represents the real values of the object's velocity. The dotting and solid lines represent the vergence velocity of the active vision system. The solid line corresponds to the vergence control with noisy conditions of the measured data ($\sigma^2 \approx 20$ pixels for the object position (u, v) in the image)

BIBLIOGRAPHY

[1] A.L.Abbot and N.A.Ahuja, "Surface reconstruction by dynamic integration of focus, camera vergence, and stereo", Int. Conf. Computer Vision, tampa, pp. 532-543, Florida, 1988

[2] Aloimonos, J., A.Bandopadhay, and I.Weiss. Active vision. In Proc., 1st Int's. Conf. on Computer Vision, pp. 34-54, June 1987. Also in Int'l.J.Computer Vision 1,4,pp. 333-356, 1988.

[3] Bajcsy,R. "Active perception", Proceedings of IEEE, 79:996-1005,1988

[4] Ballard. D., " Eye fixation and early vision: kinetic depth", ICCV, Tampa, Florida, 1988, pp 524-531

[5] Ballard, D. H., "Animate Vision", TR-329, University of Rochester, USA, Feb., 1990