2003 International Symposium on Advanced Intelligent Systems
September 25-28, 2003, Jeju, Korea

A distance perception model for AVG based on a moving camera

António Cunha*, João Barroso*, José Bulas-Cruz*, João L.Monteiro**
acunha@utad.pt, jbarroso@utad.pt, jcruz@utad.pt, Joao.Monteiro@dei.uminho.pt

* CETAV – Centro de Estudos Tecnológicos, Ambientais e da Vida
Universidade de Trás-os-Montes e Alto Douro Quinta de Prados – Apartado 1013. 5001-911 Vila Real. Portugal
** Universidade do Minho, Departamento de Electrónica Industrial, 4800 Guimarães, Portugal

***Abstract* - This paper presents a distance perception model based around a moving camera, in the context of driving a self-guidance vehicle. Aligned images, by escape points, and acquired by a moving camera, present objects at different positions depending on its relative distance to camera. The objects that are farthest from the observer (the camera) gradually lose their alignment as the distance diminishes. With the current setup, this lack of alignment is noticeable up to a distance of 10 meters. In the paper, the results of real imagery tests are presented and discussed.**

I. INTRODUCTION

In the last 20 years, important research interests appeared in the area of Intelligent Transportation Systems (ITS), with world-wide support of governmental institutions, industry and research centers. The interest on Automatic Vehicle Guidance, AVG, appears in this context, dealing with the accomplishment of tasks such as road-following, obstacle-detection and collision-prevention. Proposals have been made to solve these problems, using either passive or active sensors. It has been advocated by some research groups that the usage of passive sensors, such as cameras, has great advantages over the approaches using active sensors (such as laser-based sensors and millimeter-wave radars) [1]. Several prototypes have been developed to test and evaluate proposed solutions, [2-5]. The authors believe that machine vision will have an important contribution in solving some of these problems.

Despite the technological evolution, that brought a significant increase of calculation power and reduction of the hardware costs, we are still currently far from developing an automatic system able to satisfactorily reproduce the level of performance achieved in most AVG tasks by human beings. There is still much work to be done towards achieving a fully-autonomous AVG. On this path, we propose a distance perception model based around a moving camera, which can be applied to AVGs.

This paper is organized as follows: Section II describes the details of the distance perception model and vision algorithm, Section III presents the details of the developed prototype, Section IV presents some results, and finally Section V ends the paper with conclusions.

II. DISTANCE PERCEPTION MODEL

The distance perception model main goal is to perceive within a scene the distance to each object, by using a single camera. The basic idea is to move a camera in circular motion around a fixed axis (**e**), from position A to B (Figure 1).

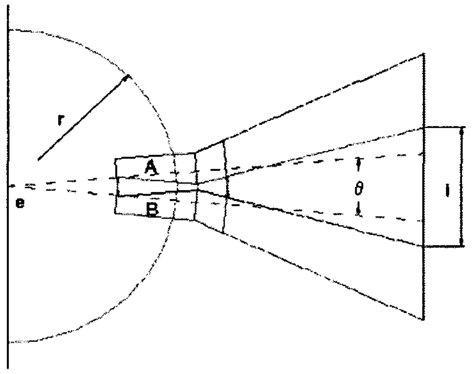

Figure 1 Circular camera motion

From this motion, two distinct images result, both of which have in common area **i** from the scene.

Increasing the amplitude of angle **θ**, the area **i** becomes increasingly thinner, which may be seen as a drawback. However, during motion, a full image sequence is acquired, with an adequate frame rate, so that the entire motion can be covered with contiguous **i**-bands.

What we get is that for a still scene (or when the camera movement is fast enough for the scene to seem still), the motion around the axis causes a change in distances.

The changes can be explained by resorting to the camera pinhole model (Figure 2). This model performs a perfect perspective projection, transforming 3-dimensional space into a 2-dimensional representation.

The projection of the perspective is characterized by the fact that all projecting lines meet at a single point C, the projection centre (C1 for camera position 1, C2 for camera position 2).

The straight line starting from point P (on the object), linking it to the projection centre C, is called the projector or projection beam; the projection centre represents, in a simplified manner, the optical centre of the objective lens.

The plane Π is the projection plane, at a distance f (focal distance) from the projection centre. Inside a real camera, this plane is the CCD, where the image is formed.

In Figure 2, planes Π_1 and Π_2, are represented for both positions of the camera, after rotating an angle θ around the **e** axis, and two points P and Q, found at different distances.

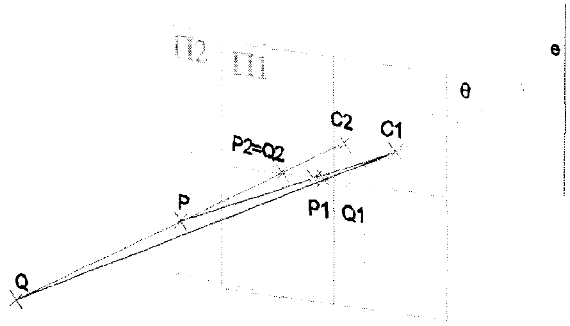

Figure 2 - Pinhole model

Projection of points P and Q, on projection plane Π_1, is given by points P1 and Q1, where the projection beams meet Π_1. Similarly, points P2 and Q2 on plane Π_2 are obtained.

We can see that the position of the projections of points P and Q varies, depending on the distance to the projection plane. The situation above has Q1 and P2 apart on the Π_1 plane and overlaid on the Π_2 plane.

If we consider that the camera in motion relatively to fixed points, we can conclude that the points nearer to the projection plane are moving faster than the ones farther away.

Should the object have a larger size (like a car, for instance) and near, another interesting effect occurs. If we split an image vertically, the points belonging to objects (the car) within the picture half to which the camera is moving in rotation, will be located at shorter distances, while the points belonging to objects within the other picture half will be at longer distances.

This change is most noticeable on objects that are closer to the observer (camera), and diminishes on objects located farther away objects.

We therefore propose to compare distinct images, resulting from simple camera rotation movements, by aligning them through points located on the horizon (the farthest away points that can be found in the images).

As a result, we get an apparent motion of objects in the image plane, which is related to the distance to the camera. The objects closer to the camera exhibit a lager apparent motion than the others.

III. DEVELOPED PROTOTYPE

In order to test the presented perception model, we developed a motorized mechanical support (Figure 3), which is able to rotate freely along the X and Y axis.

Within the several engines that could be used, we chose servo-motors, since these incorporate a stepback gearbox, with a high output torque.

We applied a position sensor at each axis (Hewlett Packard HEDS-550X/554X), allowing us to obtain an approximate resolution of 0.35 for the camera position.

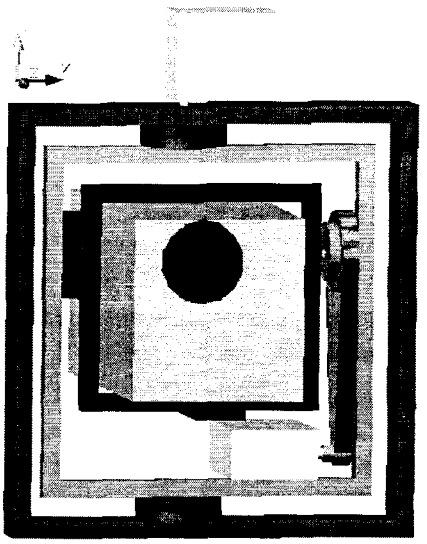

Figure 3 Motorized support for camera motion

The motorized support features an electronic control board (Figure 4) featuring two 5V power supplies (one for the logical circuitry, another for the motors), an FPGA (Xilinx XC2S100) were we implemented the logic circuits for communication, motor control and sensor reading. The interface modules perform signal adaptation between the FPGA and the exterior PC parallel port, motors, and sensors.

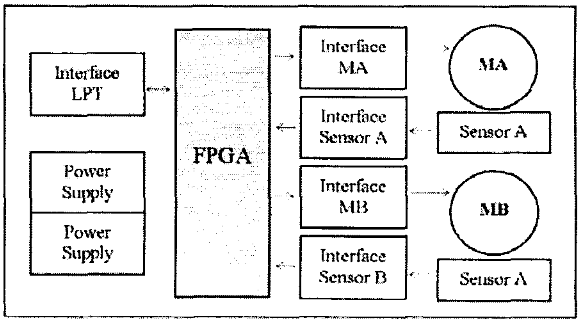

Figure 4 Electronic control board

Communication with the PC is bidirectional, through the parallel port. We chose to use an FPGA due to its enormous versatility regarding digital circuit development: an FPGA can be set-up and re-set in a matter of minutes.

The vision program generates the values for the camera positioning and transmits them to the FPGA. The FPGA upon receiving the values generates PWM signals, to control the positions of both servomotors. Finally, it answers the PC, transmitting the current values for the positions of both axes.

The computer-vision program, in the computer, controls the position of the camera, by sending to 2 BYTE values to the electronic control board. The board acts upon the servo-engines, to achieve the requested position. It also reads the two position sensors, located at the rotation axes, and sends the current readings back to the program.

After performing the initial positioning tests, we realized that the motors presented a positioning error of about 1 due to some mechanical slack.

However, due to the characteristics of the proposed model, we also realized that such errors do not prevent us from perceiving distances. The required misalignments on nearby objects, allowing us to perform distance perception using two possible positions, do not require high precision in the camera position. We must emphasize that having a more precise position tracking system, using sensors, will simplify the automatic image-alignment process, currently in its initial development phase.

The developed support is presented on Figure 5.

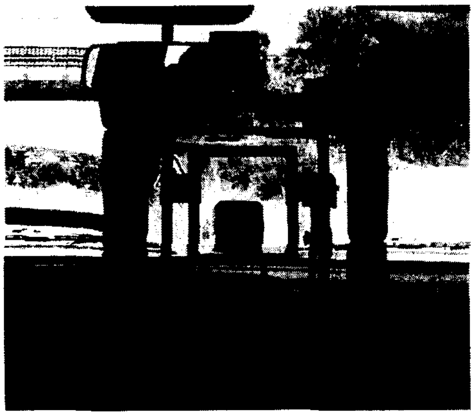

Figure 5 Developed Prototype

The camera is a SONY DFW-V500, able to acquire digital images with a resolution of 640x480 pixels, 8-bit color, at a rate of 30 fps.

The initial testing was aimed at finding a configuration that would allow us to test the proposed perception model.

By analyzing the collected images, we found that fast camera motions produce unfocused images, due to the low shutter speed of the camera. We could have used a camera with a faster shutter, but decided to reduce the camera rotation speed instead, to streamline the testing process.

We also found that to demonstrate the intended effect, we could start by using images overlapping just by $^{1}/_{3}$.

Finally, we realized that instead of having the camera hanging over the dashboard, we should place it a bit further ahead, on the vehicle front bumper, because the system is more sensitive to the first few meters; by moving the camera to the front, we can make better use of the acquired images.

IV. EXPERIMENTAL WORK

The collection of image sequences was performed with the following settings:

Camera position: vehicle front bumper.

Settings: Vehicle standing still; stationary scene.

Goal: to characterize the image impact of moving the camera between two images.

The 3 images presented in Figure 6, allow for a characterization of the proposed method.

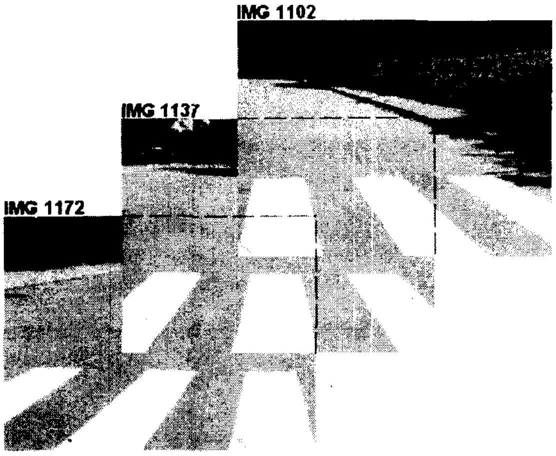

Figure 6 Image sequence gathered in the UTAD campus

The settings, in absolute standstill, were selected specifically in order to isolate the impact of the camera motion. We chose a straight line (a pedestrian crossing), 2 meters in front of the vehicle, and positioned the camera vertically, in order for the horizon and the pedestrian crossing to both be within the field of view.

The pedestrian crossing is useful, since it is a well-defined object, near the observer. The crosswalk line is a reference object that starts nearby and stretches to the horizon (approx. 80 m). Finally, in the horizon, the vegetation is used as a reference for alignment of images.

In Figure 6 we can also see a tire dirt mark, which is present in all three images.

At the gate, above the dirty stripe, we identified the position of a specific position on each image, drawing a cross over it, in order to perform the alignment.

After aligning the images, we cropped the pictures, in order to analyze only the section that was common to all 3 images. Obviously, differences are smaller between close frames, so, on Figure 7, we only present the most significant ones, frames 1172 and 1102.

We drew black lines over these images, to emphasize the contour change between both images.

The last image in Figure 7, 1172+1102, presents the sum of the two previous images, after two edge-detection processes.

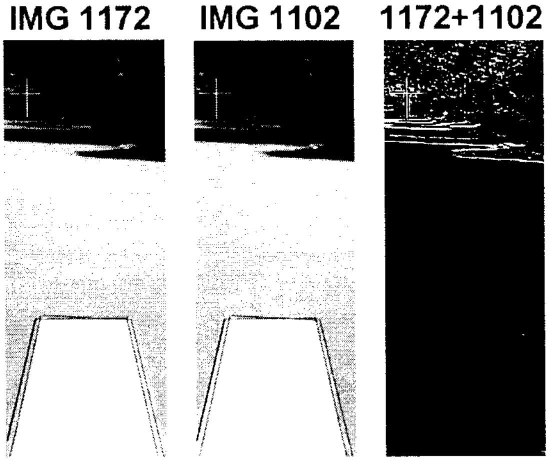

Figure 7 Mismatch between overlaid areas

Another test was performed at the same setting, but this time with a vehicle stationed near the pedestrian crossing. The alignment was made at points marked with a white cross (left of the vehicle). We present the result in Figure 8.

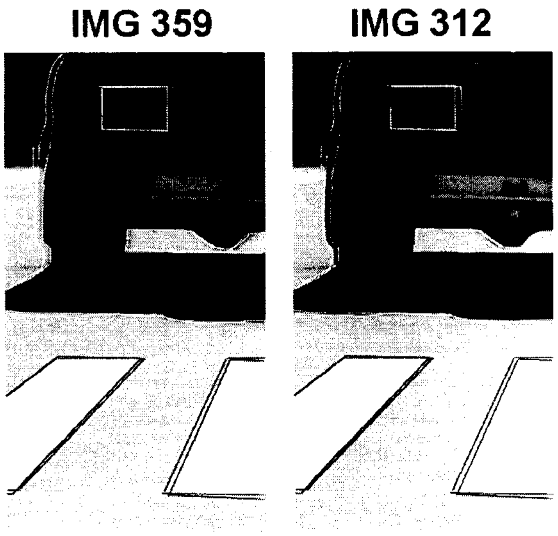

Figure 8 Mismatch between overlaid areas

As one can see, there is a noticeable deviation on the nearby objects the pedestrian crossing stripes and the vehicle. This deviation is particularly noticeable along the lines that are perpendicular to the camera motion path, i.e., the vertical lines. However, since a circular motion around an axis is taking place, we can notice that there is also a slight rotation between the two images. This is clearly visible due to the misalignment of the stripe borders, on the pedestrian crossing.

V. CONCLUSIONS

In this paper, the first version of an AVG-oriented model for perception of distances, based on a mobile camera, has been presented. This version has been used to perform preliminary tests. These tests were conducted on real-world settings, with standstill objects. With the current set-up, we managed to perceive distances up to 10 meters. We came across some problems with visual focus, causing image misalignment. Although these are preliminary results, they demonstrate the viability of the proposed model.

The authors recognize that more work is needed in order to refine the model and establish the adequate settings for different working conditions.

REFERENCES

[1] M. Bertozzi, A. Broggi, and A. Fascioli, "Vision-based Intelligent Vehicle State of the art and perspectives," presented at Robotics and Autonomous Systems, 2000.

[2] G. Ganesh Dimo, "A Panoramic Imaging System for the HSV," in *Mechanical and Mechatronic Engineering*. Sydney: University of Sydney University of Sydney, 1999, pp. 116.

[3] A. Broggi, M. Bertozzi, A. Fascioli, and G. Conte, "Automatic Vehicle Guidance: the Experience of the ARGO Autonomous Vehicle," Singapore, 1999.

[4] J. Donald, "Application of the Hough Transform to Lane Detection and Following on High Speed Roads," presented at IMVIP2001 - Irish Machine Vision & Image Processing Conference, 2001.

[5] D. Pomerleau and T. Jachem, "RALPH - Rapidly Adapting Machine Vision for Automated Vehicle Steering," in *Machine Vision*, 1996, pp. 19-27.