A Random Access 3D/2D Vision Sensor

D. Laurendeau[a], C. Gosselin[b], F. Caron[b], S. Comtois[a], T. Laliberté[b], F. Blais[c], F. Loranger[a]

[a]Computer Vision and SystemsLaboratory
Dept. of Electrical Engineering
Laval University, Ste-Foy, Quebec, Canada, G1K 7P4
[laurend,scomtois,loranger]@gel.ulaval.ca
http://www.gel/ulaval/ca/~vision

[b]Robotics Laboratory
Dept.of Mechanical Engineering
[gosselin,caron,laliberte]@gmc.ulaval.ca
http://wwwrobot.gmc.ulaval.ca

[c]Institute for Information Technology
National Research Center of Canada
M-50, Montreal Rd,Ottawa, K1A 0R6
blais@iit.nrc.ca
http://www.vit.iit.nrc.ca/~blais

Abstract

This paper presents a random access 3D/2D vision sensor. The sensor, which can be oriented in space with a fast and accurate 2 DOF mechanism, provides 2D color images as well as a single range reading at a rate of 15 Hz for each angular position of the scanning mechanism. The accuracy of the range readings is 0.8% at 1.5m.

Résumé

Cet article présente un capteur de vision 3D/2D à accès aléatoire. Le capteur, qui peut être orienté rapidement et avec précision grâce à un mécanisme à 2 degrés de liberté en rotation, fournit des images 2D couleurs de même qu'une donnée de profondeur à une cadence de 15 Hz pour chaque position angulaire du mécanisme de balayage. L'exactitude des lectures de profondeur est de 0.8% à 1.5m.

1. Introduction

In active vision and robotics applications, there is a need for vision sensors that can focus their attention on specific parts of the environment in an efficient and flexible manner. Furthermore, such sensors should provide registered 2D and 3D images of the environment, thus allowing sensor fusion between the two sensing modalities for scene modeling, path planning, and collision avoidance.

The advent of the World Wide Web has also raised the issue of flexible vision sensors. Applications are now envisioned for which virtual environments will be built from range and color images of actual scenes. It will then be possible to fly through these environments using VRML-capable browsers. In these applications, the vision sensor must provide registered 2D and 3D images of the scenes and must offer gazing control for an efficient yet exhaustive sampling of the scene.

This paper presents an active vision sensor that allows the acquisition of registered 2D color images and 3D range data. The sensor is mounted in a mechanism that allows to orient the sensor in any direction of the field of view, making it an attractive device for active vision, focus of attention, and gazing experiments. It is also well adapted for the construction of virtual models of existing sites since it allows flexible scanning of the environment.

The paper is divided as follows. Section 2 describes the architecture of the vision sensor and the approach that was chosen to acquire 2D color images and 3D range data. The performance of the sensor are given in terms of speed and accuracy. Section 3 describes the 2 degrees of freedom (DOF) mechanism that allows to orient the vision sensor and to scan the 3D data in almost any trajectory. Section 4 presents experimental results of the acquisition of 3D and 2D images and demonstrates the flexibility of the active vision sensor. Section 5 concludes and describes future applications of the sensor.

2. The 2D/3D vision sensor (TRID)

This section describes the architecture of the sensor and the 2D and 3D image acquisition approach. In the following, the sensor is called *TRID* for *TRI*angulation-*D*efocusing sensor.

2.1 Basic components of the sensor

Figure 1 (a) shows the vision sensor. The dual vision sensor is composed of the following components:

1- a miniature color CCD camera (Panasonic KS 202) for the acquisition of photometric data on the scene;
2- a miniature laser diode (1 mW, λ ~ 630 nm, Class II) with focused beam;
3- a b/w CCD camera (Panasonic WD-CD50) with a 50 mm lens. A mask with a circular opening is placed in front of the lens. The annular mask is shown in Figure 1 (b).

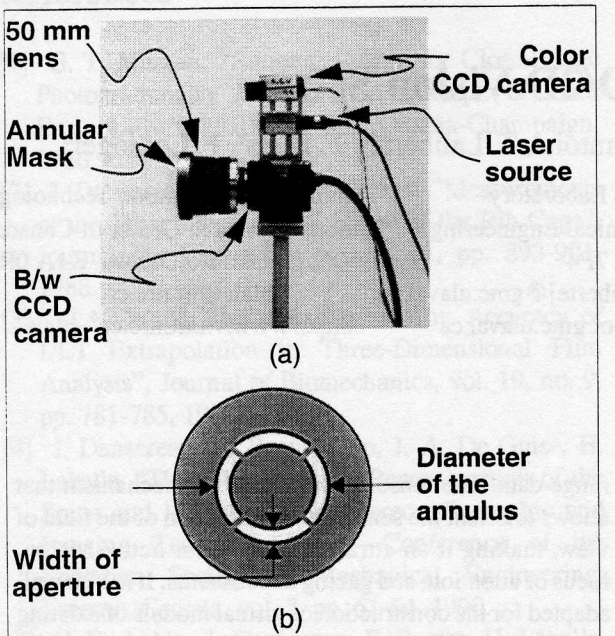

Figure 1 . The vision sensor. (a) The cameras & laser, (b) the annular mask.

2.2 Range measurement approach

The measurement of range data is performed on a single point in the scene and combines a range from defocusing approach (i.e. a version of the Biris sensor [1]) with an active triangulation approach, both implemented on the same camera. Figure 2 shows the geometry for range measurement. A laser beam is directed toward the scene. The laser spot reflected on the surface is imaged on a defocused camera through the circular mask shown in Fig. 1 (b). The image of the spot is a circle of radius D. Based on the Biris equations (see [2]), the range z_b between the sensor and the surface is related to D:

$$z_B = \frac{D\,s(l-f)+dfl}{D(l-f)+df} \qquad \text{Eq.1}$$

where the parameters of Eq.1 are described in Table 1.

Table 1: Parameter definitions for Eq.1

Param.	Description
D	diameter of the circle
s	distance between the lens and the mask
l	distance between the reference plane and the focal point of the lens
f	focal length of the lens
d	diameter of the aperture on the mask
z_B	estimated z coordinate of the laser spot using the Biris approach

If the position of the center of gravity of the circle is computed, the geometric arrangement in Figure 2 can also be used as a classical triangulation range finder. In this case, range z_T is given by:

$$z_T = \frac{\Delta f \tan(\theta)}{\rho_y \tan(\theta) - f} \qquad \text{Eq.2}$$

where the parameters in Eq.2 are described in Table 2.

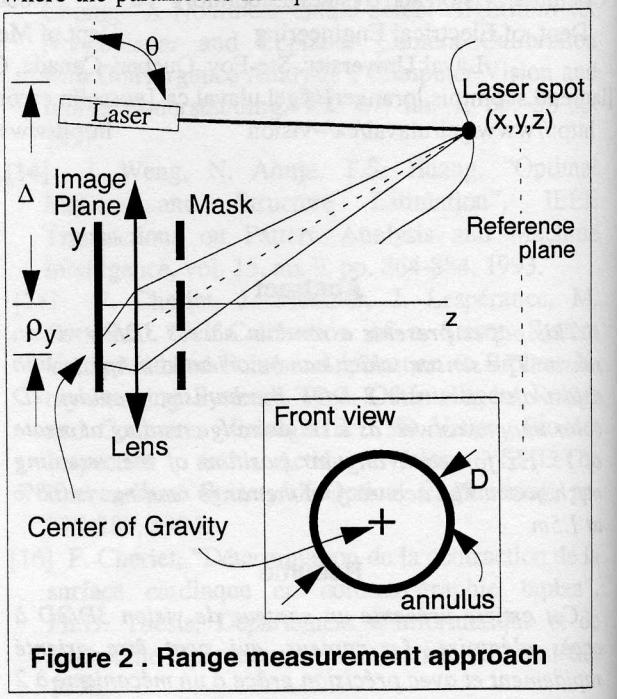

Figure 2 . Range measurement approach

The x and y coordinates are computed as follows:

$$y = \frac{\rho_y \Delta f \tan(\theta)}{f(\rho_y \tan(\theta) - f)} \qquad \text{Eq.3}$$

$$x = 0 \qquad \text{Eq.4}$$

where ρ_y is the y coordinate of the center of gravity of the image of the laser spot as seen through the annular mask. Since the position of the laser changes only along the Y axis (e.g. the laser travels along the Y axis at $\rho_x = 0$, the x coordinate of the 3-D points is always 0).

Table 2: Parameter definitions for Eq.2 - Eq.4

Param.	Description
Δ	Baseline (distance between the optical center and the laser)
f	focal length of the lens
θ	angle between the laser beam and the image plane
ρ_y	distance between the center of gravity of the circle and the optical center on the image plane
z_T	estimated z coordinate of the laser spot using triangulation

The two measurements z_B and z_T are combined

using a Kalman filtering approach:

$$z_K = az_T + (1-a)z_B \qquad \text{Eq.5}$$

where:

$$a = \frac{\sigma_B^2}{\sigma_T^2 + \sigma_B^2}. \qquad \text{Eq.6}$$

The variance of z_K is given by:

$$\sigma_K^2 = a^2\sigma_T^2 + (1-a)^2\sigma_B^2. \qquad \text{Eq.7}$$

The combination of both approaches has several advantages over each method taken separately:

1- the statistics for computing the center of gravity of the circle is much larger than the one that would be available if the center of gravity of the focused image of the laser spot would be computed. The estimate of the image point is thus more accurate;

2- the use of Biris allows to decrease the baseline that would be required for the classical triangulation to give identical accuracy alone.

The main disadvantage of the approach is that the image processing that is required for extracting the image of the circle and for computing the position of its center of gravity is more complex than for the classical active triangulation.

2.3 Image processing for range extraction

The image processing that is required for the estimation of range is simple but must meet a real-time constraint. Fig. 3 shows the image of the defocused circle as seen through the annular mask (in reverse video for clarity). The processing of this image must yield the following parameters:

1- the diameter D of the circle used in Eq.1;

2- the coordinates (ρ_x, ρ_y) of the center of gravity of the circle used in Eqs 2, 3 and 4[1].

The processing has been divided into three steps:

1- the estimation of the region circumscribing the circle;

2- the detection of the pixels belonging to the circle;

3- the estimation of the parameters of the circle that best describe the circle.

These steps are based on the approach reported in [3].

2.3.1 Estimation of the region circumscribing the circle

In order to estimate this region, the grey-level histogram of the image is computed. This histogram is computed only at start-up time or when the region cannot be estimated satisfactorily (for instance when the return signal from the laser source is too low). The histogram is analyzed in order to find a threshold for separating the circle from the background. The value of the threshold is chosen such as at least 500 pixels are above this value. This value has been chosen based on experiments in different lighting conditions and with objects of different surface texture (white matte plane, dark wooden beam, grey cardboard, etc...).

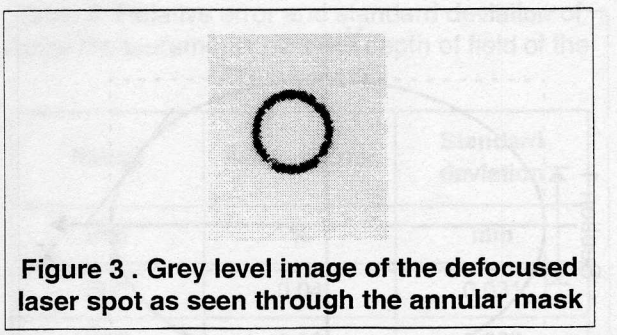

Figure 3 . Grey level image of the defocused laser spot as seen through the annular mask

The image of the circle is thresholded and only the pixels with a grey level above the threshold are considered for further processing. At start-up, the thresholding is performed on the whole image. In steady state mode, it is only performed in an region where the circle was previously found. If the circle is not found in this region, the area of the region is gradually expanded until the circle is found or until it is required to recompute the histogram. Once detected, the area of the region is checked to find whether it can be associated with the circle or not. Since the range finder has a limited depth of field ($0.5m <$ Depth of Field $< 1.5m$), it is possible to know in advance the maximum and minimum values of the diameter of the circle and thus to use these values for guiding the image processing algorithm.

2.3.2 Detection of the pixels belonging to the circle

Once the region containing the pixels belonging to the circle has been detected, the next step is to find the pixels along the circle. This is accomplished by finding the zero crossings of the first derivative of the image along the perpendicular to the tangent to the circle. In practice, to speed up the computations, the derivative is computed horizontally for section 1 of the circumscribing rectangle and vertically for section 2, as shown in Figure 4.

The filter that is used is a low-pass derivative adaptive filter of the form [1 1 1 ... 1 0 -1 ... -1 -1]. The length of the filter is chosen according to the width (height) of the region circumscribing the circle. This adaptive behavior reduces the speckle effect of the laser.

The computation of the second derivative allows to validate the zero crossings of the first derivative. Only the pixels with a value of the second derivative that

1. Actually, $\rho_x = 0$ as mentioned in Section 2.2

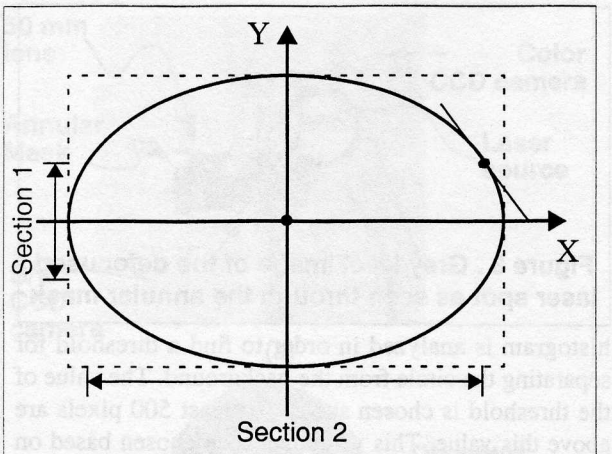

Figure 4 .Sections for computing the first derivative

exceeds a given threshold are considered for further processing[2]. The threshold is chosen adaptively in order to deal with signals of varying amplitude. To speed up the computation, the second derivative is computed only in the vicinity of the zero crossings of the first derivative. Figure 5 shows the result of this processing step for a single line of the video signal. Only the zero crossings of the first derivative that are validated by a high value of the second derivative are labeled as belonging to the circle.

2.3.3 Estimation of the parameters of the circle

A least-squares algorithm fits a circle to the pixels that remain after the processing steps described in Sections 2.3.1 and 2.3.2 and the parameters of the circle (D and (ρ_x, ρ_y) can then be used to compute z_K, x, and y.

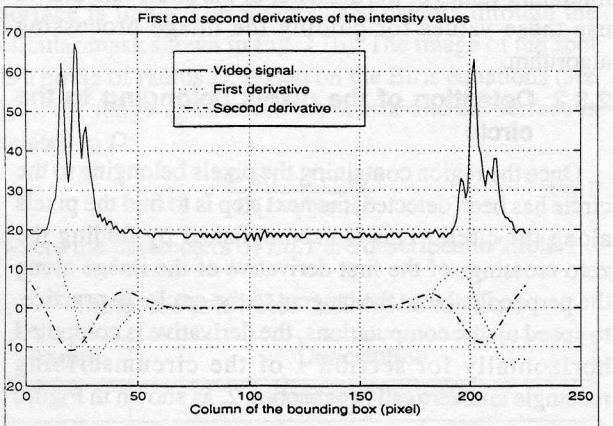

Figure 5 .Example of the detection of relevant pixels on the circle for a single video line.

2.4 Calibration of the sensor

Equations 1, 2, 3, 4 describe the behavior of the sensor for the estimation of 3-D coordinates. However, in order to take into account several phenomena that are not modeled by the equations[3], these 3-D coordinates are rather obtained by look-up tables built from a calibration of the sensor. The different steps of the calibration procedure are:

1- choosing of the baseline Δ and the direction angle of the laser beam θ (see Figure 2);

2- construction of the look-up tables for the measurement of range values;

3- estimation of the variance of the estimated range values provided by the range from defocusing and the active triangulation approaches.

2.4.1 Choosing the baseline and direction angle of the laser

The desired work interval of the sensor is between 0.5 m and 1.5 m. Although the baseline and direction angle of the laser beam can be computed from the geometry of the sensor and the CCD, they are are adjusted manually such that the vertical position of the center of gravity of the circle in the image (i.e. σ_y) covers the whole field of view over this range interval. Figure 6 shows this procedure.

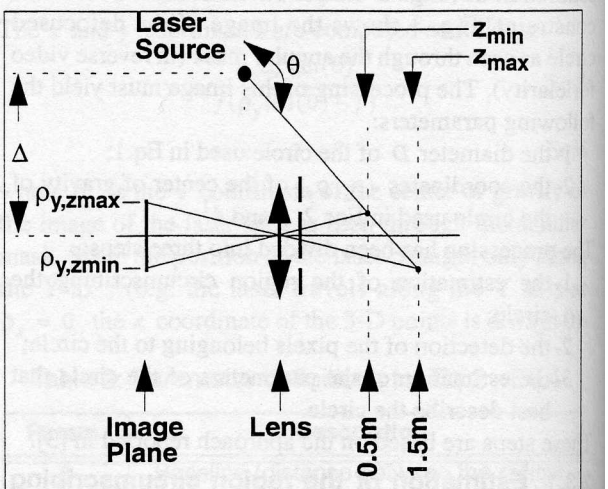

Figure 6 .Setting of the baseline and direction angle of the laser

2.4.2 Construction of the look-up tables for the measurement of range values

The look-up tables are built by measuring with the

2. Actually, the image of the laser spot is not a circle but rather an ellipse. This results from the fact that the horizontal and vertical scale factors of the camera are different. The ellipse is transformed into a circle by a proper calibration of the scale factors. See [3] for details.

3. Lens distortion, etc.

sensor the range values z_B and z_T of a plane that is moved at 101 different positions on an optical bench between 0.5 m and 1.5 m. For each position of the plane, 30 images are taken and the average value of D, and σ_y as well as their variance are computed. At the end of the calibration procedure, 4 look-up tables are available:

$$\Delta = f(z) \quad \text{Eq.8}$$

$$\sigma_\Delta^2 = f(z) \quad \text{Eq.9}$$

$$\rho_y = f(z) \quad \text{Eq.10}$$

$$\sigma_{\rho_y}^2 = f(z) \quad \text{Eq.11}$$

These tables are linearly interpolated in order to obtain a constant step relative to the parameters (Δ, ρ_y) instead of the range value (i.e. z_B and z_T becomes a function of Δ).

The Kalman filtering equation 5 requires the knowledge of the variance of both range estimates z_B and z_T. The variance of parameters Δ and ρ_y are available for the look-up tables (Eqs 8 and 9). Even though the z_B and z_T are not linear functions of Δ and ρ_y (see Eqs 1 and 2), their variance can be computed with an error propagation method. It is then possible to compute two additional look-up tables:

$$\sigma_{z_B}^2 = f(z_B) \quad \text{Eq.12}$$

$$\sigma_{\rho_y}^2 = f(\rho_y). \quad \text{Eq.13}$$

The intrinsic and extrinsic camera parameters for the color CCD camera are obtained using the Tsai algorithm [4].

2.5 Performance of the sensor

Several experiments were conducted with the sensor in order to evaluate its performance under variable conditions. The most interesting of these experiments is the characterization of the accuracy of the sensor for range measurements. The result of this experiment is reported in Table 3.

Table 3: Relative error and standard deviation of range measurements over the depth of field of the TRID sensor.

Range	Relative Error	Standard deviation
mm	%	mm
500	0.05	0.097
600	0.01	0.003
700	0.02	0.023
800	0.04	0.031
900	0.06	0.085
1000	0.11	0.088
1100	0.18	0.049
1200	0.30	0.056
1300	0.42	0.097
1400	0.52	0.133
1500	0.80	0.143

These results show that the standard deviation of the range estimates is less than 0.2mm at 1.5m. This standard deviation decreases as the object gets closer to the sensor.

2.6 Hardware implementation of TRID

The TRID sensor image processing algorithms described in Section 2.3 were implemented on Spectrum TMS320C40-based processing boards running at 40 MHz. The video frame grabber (a Spectrum MDC40IM image grabber) resides on a first board and is also responsible for the communications with the host computer (a Sun Workstation. See Section 3.3), with the image processing board for TRID, and with the control board for the scanning mechanism (See Section 3.2).

With the current hardware implementation, TRID provides a range measuring rate of 5Hz at 0.5m and ~15 Hz at 1.5m. Remember that the diameter of the circle increases as range decreases and thus requires more processing time for extracting the pixels belonging to the circle.

3. The scanning mechanism

The TRID Sensor provides a single range reading per frame. A 2 degree of freedom (DOF) mechanism has been designed to allow the acquisition of both random and structured patterns of range data by orienting the sensor in the area of interest. Figure 7 shows the basic principle of the mechanism (a), and a photograph of the prototype that has been implemented in our lab (b). With this mechanism, it is possible to orient the vision sensor in a wide area of the scene and to acquire 3-D (and 2D)

data along any possible path.

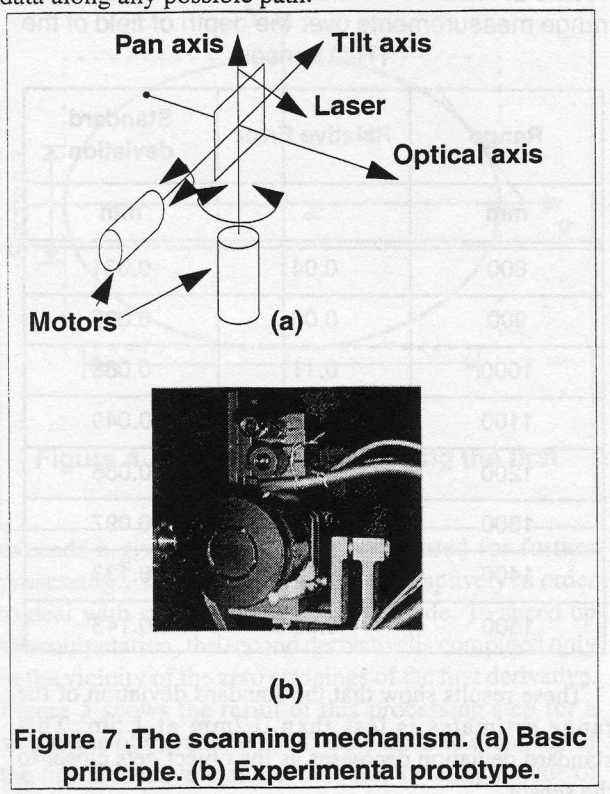

Figure 7 .The scanning mechanism. (a) Basic principle. (b) Experimental prototype.

The specs of the 2 DOF mechanism are given in Table 4. The speed and precision of the mechanism offer a broad spectrum of applications ranging form tracking of moving targets to the fast scanning of a scene with a programmable path.

Table 4: Specifications of the 2 DOF mechanism

Parameter		Value	Units
workspace			
	minimum pan angle	-45	deg
	maximum pan angle	+45	deg
	minimum tilt angle	+45	deg
	maximum tilt angle	-45	deg
resolution			
	pan angle	0.18	deg
	tilt angle	0.18	deg
precision			
	pan angle	±0.09	deg
	tilt angle	±0.09	deg
maximum angular speed			
	pan angle	2666	rpm
	tilt angle	1333	rpm

3.1 Control of the mechanism

Figure 8 shows the architecture of the control unit of the 2 DOF mechanism. The user input consists in the desired position, speed and acceleration for the mechanism at a given time.

The interpolating polynomial computes two triplets of values (P,V,A) in cartesian coordinates, one for each axis (pan and tilt), in response to the user input. These triplets are computed at each control cycle (0.001 s). The static and dynamic models convert the position values in cartesian coordinates into their articular correspondence. They also estimate the input that must be fed to the motors, given the desired position, speed and acceleration. The purpose of the controller (see Figure 8) is to compensate for the modeling errors in the static and dynamic models of each axis. Its role is to maintain the position of the mechanism at the position given by both the static and dynamic models.

3.2 Electronics of the controller

The control software for the system described in Section 3.1 above has been implemented on a TMS320C40 DSP board that belongs to the same unit as the frame grabber and the image processing board for TRID algorithms. A PCB has been designed for interfacing the power amplifiers and the I/O board, for implementing an electronic watchdog, and for controlling the brakes on the stepping motors that drive the axes of the mechanism.

3.3 Graphics User Interface

The use of the random access range finder is made easy through the graphics user interface (GUI) that has been placed on top of the TRID and controller software modules. The GUI runs on a Sun SPARC workstation. From the GUI, it is possible to:

1- reset and calibrate the system;
2- download the TRID software and calibration tables from the SPARC workstation onto the TMS320C40 board;
3- download the control software from the SPARC workstation onto the TMS320C40 board for the 2 DOF mechanism;

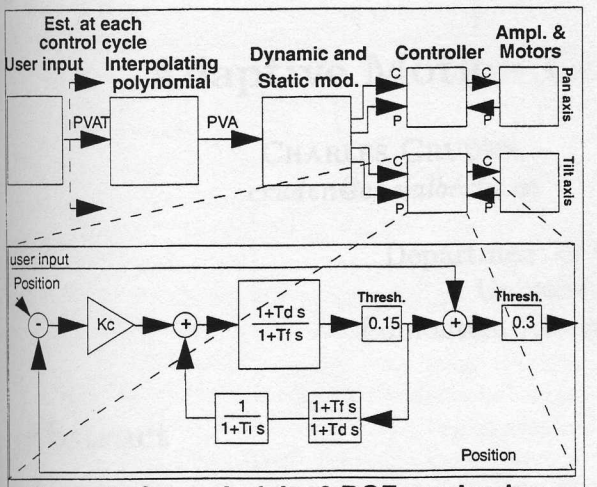

Figure 8 .Control of the 2 DOF mechanism

4- program a trajectory for scanning a scene;

5- monitor the status of the range sensor and the 2 DOF mechanism;

6- transfer computed range data to the host computer;

It is important to mention that is possible to acquire any number of range samples for a given pan and tilt combination. It is thus possible to acquire 3-D images rapidly by sampling a single range reading at each pan/ tilt position or to acquire very accurate images by averaging several range readings at each position.

4. Experimental results

To demonstrate the performance of the scanning mechanism combined with the TRID sensor, we have scanned the pyramid in Fig. 9 (a). The pyramid is seen from the point of view of the color camera shown in Fig. 1. The laser spot used by TRID is also visible near the front edge of the pyramid.

The pyramid scene was scanned along a spiral pattern (Fig. 9 (b)) and along a raster pattern (Fig. 9 (c)).

5. Conclusion

We have presented a random access range sensor composed of a dual mode range finding camera (defocusing and active triangulation) and a 2 DOF rotation mechanism. The range finder provides an 0.8% error on single range measurements at 1.5 m at a rate of 15 Hz. The 2 DOF mechanism allows to orient the dual sensor at very high speed and with very good angular accuracy. The random access range finder image processing algorithms and controller software are implemented in C language and run on TMS320C40 DSP boards.

Future work will focus on the fine tuning of the controller as well as on the implementation of the system on a Pentium-based architecture running under QNX.

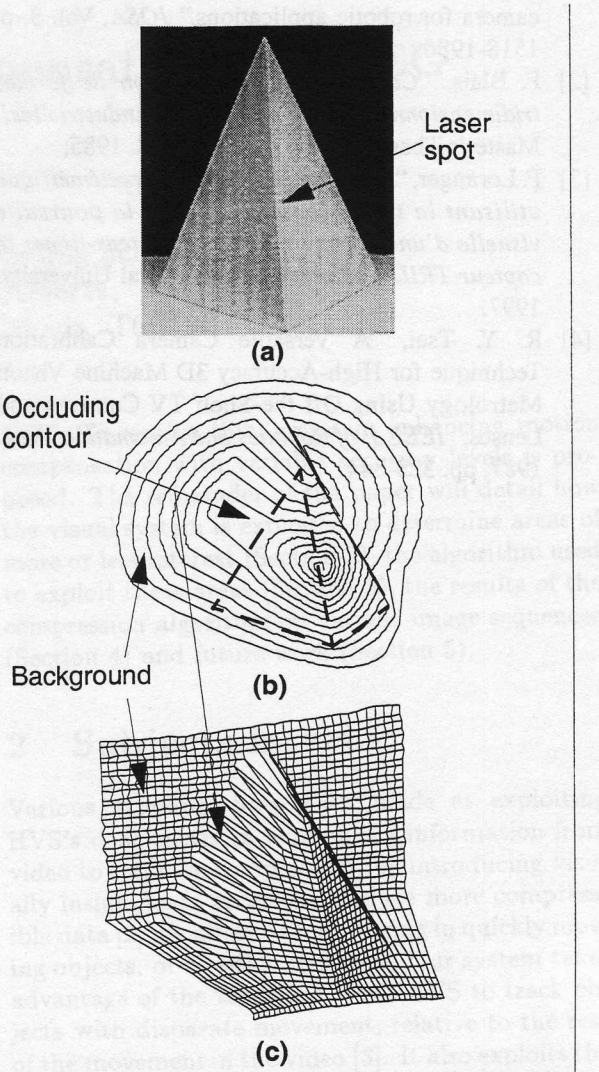

Figure 9 .Experimental results. (a) Pyramid scene. (b) Spiral pattern. (c) Raster pattern.

6. Acknowledgments

DL, CG, SC, and FL are members of the Institute for Robotics and Intelligent Systems (IRIS) and wish to acknowledge the support of the Networks of Centres of Excellence Program of the Government of Canada, the Natural Sciences and Engineering Council, and the participation of PRECARN Associates Inc.

The research was supported by a contract with the Hydro-Quebec Research Institute (IREQ, Varennes, P.Q.).

DL, SC and TL would like to thank Dr André Desbiens from Laval University for his help in the design of the controller.

7. References

[1] M. Rioux, F. Blais, "Compact three-dimensional

camera for robotic applications," *JOSA*, Vol. 3, p. 1518-1986

[2] F. Blais, "*Capteur optique de vision de formes tridimensionnelles pour applications industrielles*," Master's Thesis, Laval University, oct. 1985.

[3] F. Loranger, "*Réalisation d'un capteur télémétrique utilisant la technologie Biris pour la poursuite visuelle d'une structure poutre-isolateur-ligne: le capteur TRID*," Master's Thesis, Laval University, 1997.

[4] R. Y. Tsai, "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses," *IEEE J. of Robotics and Automation*, Aug. 1987, pp. 323-344.