Laser Eye – a new 3D sensor for active vision

Piotr Jasiobedzki[1], Michael Jenkin[2], Evangelos Milios[2],
Brian Down[1], John Tsotsos[1], Todd Campbell[3]

[1] Dept. of Computer Science, University of Toronto
Toronto, Canada M5S 1A4

[2] Dept. of Computer Science, York University
North York, Canada M3J 1P3

[3] Atomic Energy of Canada Ltd.
2251 Speakman Dr., Mississauga, Canada L5K 1B2

ABSTRACT

This paper describes a new sensor that combines visual information from a CCD camera with sparse distance measurements from an infra–red laser range–finder. The camera and the range–finder are coupled together in such a way that their optical axes are parallel. A mirror with a different reflectivity for visible and for infra–red light is used to ensure collinearity of effective optical axes of the camera lens and the range–finder. The range is measured for an object in the centre of the camera field of view. The Laser Eye is mounted on a robotic head and is used in an active vision system for an autonomous mobile robot (called ARK).

1. INTRODUCTION

Three dimensional information is required in many tasks performed by autonomous robots. Position estimation, navigation, obstacle avoidance, object detection and recognition all require some form of 3D data. Many computer vision tasks involving natural scenes such as image segmentation, object recognition can be improved if 3D data is available. Position estimation of an autonomous mobile robot in a known environment is performed by comparing the position of detected objects relative to the robot with their positions in a map of the environment[7, 8]. Mobile robots operating in industrial environments such as industrial bays and power stations should be equipped with range sensors capable of operating at distances of several tens of meters. Navigation and obstacle avoidance tasks require information on distance to objects in the vicinity of the robot.

Range information can be obtained by analyzing pairs of stereo images, by projecting a structured pattern of light on the scene and recovering depth by triangulation[11], by using a special iris[2] or by using scanning time–of–flight 3D sensors[5]. The accuracy and effective range of stereo and triangulation methods depend on the distance between cameras (stereo) or a camera and a projector (triangulation). 3D scanning sensors based on the time–of–flight principle can operate at much larger distances (limits are posed by the light source energy, reflectivity of the surface, etc.).

2. COMBINING DEPTH AND INTENSITY INFORMATION

Analysis of stereo images provides information on depth in addition to the grey scale or colour data present in the original images. Stereo image pairs can be obtained using, for example, a binocular robotic head (for references see[1]). The disparity measured between the corresponding features in both images, together with the known length of the baseline (distance between cameras) and the vergence angle, allows estimation of the distance to the feature.

The triangulation method uses a structured light projector and a camera. Position of a projected point (a stripe or a pattern) in the camera field of view, together with information on the distance between the camera and the projector

 0-8194-1324-0/93/$6.00

and the angle between them allows recovery of the depth to the illuminated object[10]. A combination of a binocular stereo head and a triangulation depth sensor has also been investigated[11]. Image segmentation is performed by fusing data from two intensity images and the depth image. The range of stereo and triangulation techniques depend on the baseline length, sensitivity of the method used to measure disparity and accuracy of the vergence mechanisms. In a typical robotic application with a baseline of approximately 30 cm, the depth range is up to several meters.

Laser scanners can operate at distances of up to several tens of meters. The laser scanner sweeps a laser beam across the field of view. Objects within the field of view reflect the light back to the scanner. Time of flight and intensity of the reflected light are recorded for each orientation of the scanning beam and are stored as two registered images: depth and intensity[5]. The intensity image, created by the scanner, is not a visual image as is, for example, the image received from a video camera. Laser scanners create intensity images of the reflected laser beam. Such a reflectance image contains information about properties of the materials constituting the observed objects, but it is significantly different from an optical image created under the original illumination. The cost of a laser scanner is high due to the required precision and speed of the scanning mechanism.

Fusion of data from independent depth and optical image sensors is possible and was described in the literature. The main problems here seem to be the spatial separation between sensors causing a different field of view for each sensor and that the sensors often have different resolutions. It is necessary to find correspondences between both images before the data can be fused or to transform the data from one coordinate system, associated with the first sensor, to the coordinate system of the other. Such a transformation requires accurate calibration of both instruments.

Another approach is based on using one of the sensors (video camera) as the main source of information and the other (depth sensor) to perform close inspection of selected objects in the scene. Selection of such objects is performed by analyzing data from the camera. Correspondence between the image observed by the camera and the depth measurements can be achieved easily if the inspected object is illuminated by a laser beam visible to the camera[4, 8]. The camera and the laser range-finder can be mounted and controlled independently[4] or integrated into one unit[8]. In the latter case the optical axes of the camera and laser are parallel. The unit can be aimed at an object by comparing the position of the illuminated spot to the object position or by bringing the image of the object into the centre of the camera field of view. Accuracy is limited by the parallax error between the optical axes.

3. LASER EYE

In our design we use a sensor consisting of a camera and a laser range-finder. The range-finder uses the time-of-flight principle and provides a single depth measurement for each orientation of the sensor. Measuring the distance requires pointing the sensor at an object of interest and reading the depth value. The sensor is mounted on a robotic head and can be oriented in any direction within it's range.

In the first version of our robotic head we used a camera and a laser range-finder mounted together on a pan and tilt unit[6, 8]. Optical axes of both instruments were aligned to be parallel and, in the two models which have been built, the distance between the axes was approximately 10 cm. The range-finder measured distance to a point within the camera field of view (figure 1) slightly offset from the centre of the field. The range-finder used in the design is a commercial device[9]. It can provide measurements with an accuracy better than 5 cm at distances from 0.2 m to 100 m. The range-finder uses an infra read laser diode to generate a sequence of optical pulses that are reflected from a target. The time required to travel to and from the target is measured. It is possible to operate the range-finder in a fast or slow mode and increase the resolution by averaging more pulses.

3.1. Principle of operation

The offset between the optical axes in the first version caused a parallax error[6, 8]. This error was constant in physical units but it varied in image coordinates depending on the distance to an observed object. This error did not play much of a role for wide angle lenses but became a limiting factor when we used narrow angle lenses.

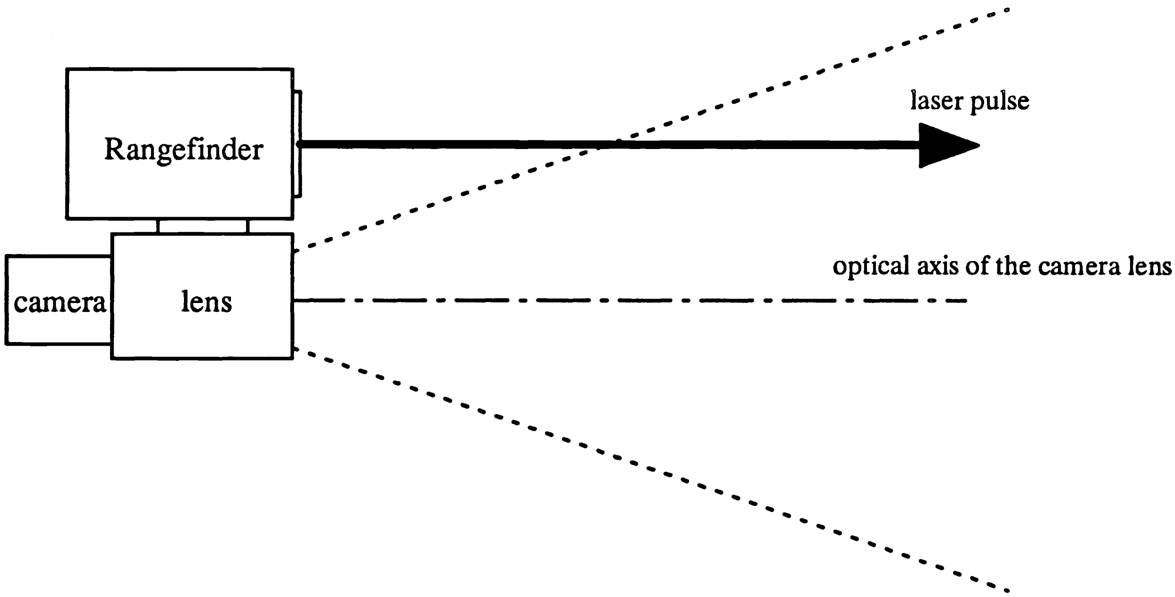

Figure 1. First version of the combined range and visual sensor (with a parallax error).

The new version of the sensor solves the problem of the parallax error by ensuring collinearity between optical axes of both instruments. The distance measured by the range–finder always corresponds to the depth of an object in the centre of the camera's field of view. The collinearity is achieved by using mirrors of different reflective properties for visible and infra–red light. Two such mirrors are available: a cold mirror (reflects visible and transmits infra–red light) and a hot mirror (reflects infra–red and transmits visible light). Version 2a of the Laser Eye uses the cold mirror (figure 2), version 2b uses the hot mirror (figure3).

The original design of the sensor (figure 1) has been modified by addition of two mirrors: a cold and a standard mirror (figure 2). The infrared laser pulse sent by the transmitter passes with minimum attenuation through the cold mirror. The reflected pulse returns to the detector through the same mirror. The visible light is reflected first by the cold mirror, then by the standard mirror and projected through the lens on a CCD element in the camera. Both mirrors are oriented in a such way that optical axis of the range–finder and the effective optical axis of the lens after reflection are collinear.

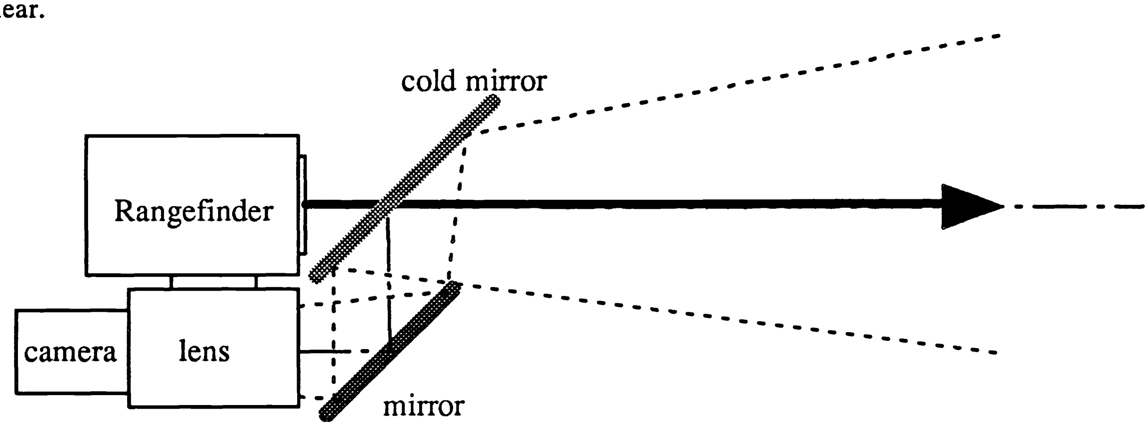

Figure 2. Modified combined range and visual sensor with cold mirror (no parallax error).

Another configuration of the Laser Eye sensor is possible if a hot and a standard mirror are used (figure 3). The hot mirror is placed in the path of light projecting into the camera lens. The visible light is transmitted through this mirror with minimum attenuation. The infrared laser pulse sent by the transmitter is first reflected by the standard mirror and later by the hot mirror. The reflected pulse returns to the detector along the same path. Again, both mirrors are oriented in a such way that the effective optical axis of the range–finder after reflection and the optical axis of the lens are collinear.

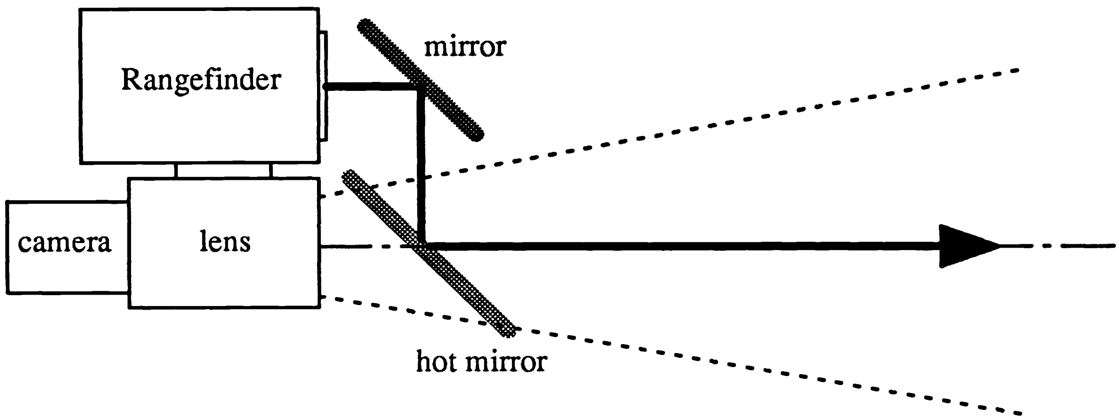

Figure 3. Modified combined range and visual sensor with hot mirror (no parallax error).

We also considered designs of similar sensors with no parallax errors that involved using of a semi–transparent, a mobile or a mirror with an aperture. The designs using hot or cold mirrors proved to be superior.

3.2. Sensor design

The Laser Eye is mounted on a computer controlled robotic head (figure 4). The robotic head has five degrees of freedom, namely pan or horizontal motion, tilt or vertical motion and camera zoom, focus and iris. The panning range of the head covers 360 degrees. The head can tilt in any direction between −65 degrees and 95 degrees above the horizon. The head can rotate with speeds up to 180 degrees / second.

In our design of the head, we wanted to use a motorised zoom lens that would allow us to obtain both a fairly wide view of a scene and a narrow field of view as well. The design utilizing the hot mirror requires smaller mirrors than the design with the cold mirror, so we decided to build a sensor based on the first concept. The range–finder is mounted above the tilt axis of the sensor and has two lenses: one for the transmitter and second for the receiver. A camera equipped with a motorized lens is mounted below the tilt axis. Two mirrors are placed in front of the range–finder and the camera lens as shown in figure 3 and explained in section 3.1.

The effective range of the range–finder used is approximately 0.2 m – 100 m and the laser beam divergence is less than 5 mrad[9]. This corresponds to a laser spot of 3 pixels in diameter for an image digitised in a 512 x 512 grid and for the wide setting of the zoom lens (45 degrees). For the narrow setting of the zoom lens (4.5 degrees), the spot is 30 pixels in diameter.

Commercially available hot mirrors ensure that, on average, 90% of the infra–red light is reflected and 85% of the visible light is transmitted[3].

4. AIMING THE SENSOR

The active spot of the range–finder is located exactly in the centre of the camera's field of view. In order to measure distance to an object or a feature of interest it is necessary to bring it to the centre. This is achieved by measuring the

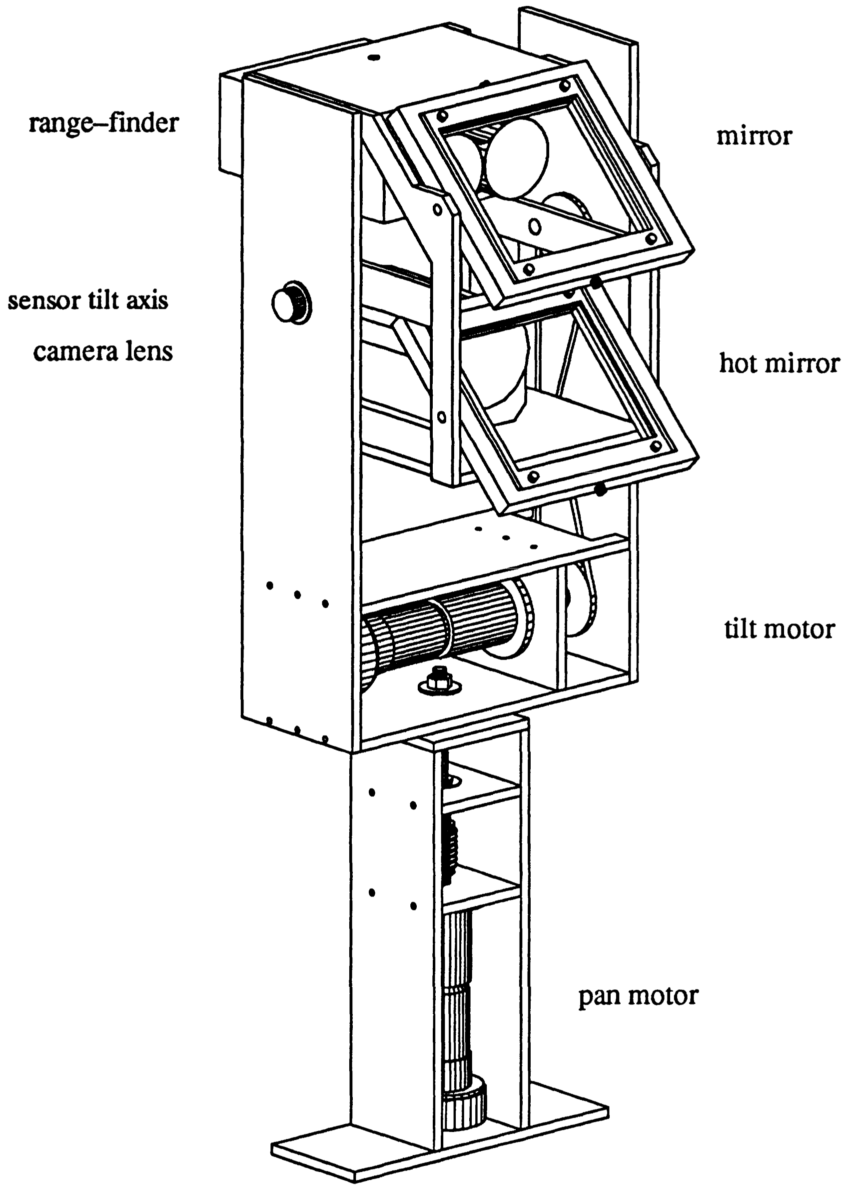

Figure 4. Robotic head with the Laser Eye sensor

displacement of the object from the centre of the initial image in image coordinates. Then the image based coordinates are converted into head coordinates using parameters of the camera, the lens and location of the lens with respect to the head pan and tilt axes. This data is sent to the head controller and used to drive the pan and tilt motors.

Higher accuracy in aiming the head requires the use of visual information in the head position feedback loop. A window around an object of interest in the initial image is selected and the head is moved in a such way so as to bring the window into the centre of the field of view of the camera. The window's offset from the centre is estimated by matching the image content of the window to new images and detecting the location of the best fit. The matching is achieved by correlation of the contents of the window with the new images. The visual based aiming of the head will be used to refine the initial position estimate obtained from head calibration data.

Each range measurement can take from 0.2 to 0.5 sec depending on the required accuracy. The time required to point the sensor in a new direction depends on the angle and is in the range 0.2 – 1 sec. The head with the Laser Eye can operate at rates of up to several measurements per second without using visual feedback.

5. CONCLUSIONS

A new a sensor combining visual information with sparse range measurements has been presented. The sensor consists of a camera, an infra–red laser range–finder and a mirror of different reflectivity for visible and infra red light. The range–finder measures distance to an object in the centre of a camera field of view. The sensor is mounted on a robotic head with pan and tilt capabilities. The main application is in an active vision system for an autonomous mobile robot.

6. ACKNOWLEDGEMENTS

Funding for this work was provided, in part, by the ARK (Autonomous Robot for a Known environment) Project, which receives its funding from PRECARN Associates Inc., the Department of Industry, Science and Technology Canada, the National Research Council of Canada, Technology Ontario, Ontario Hydro, and Atomic Energy of Canada Limited.

7. REFERENCES

1. *Applications of Artificial Intelligence X: Machine Vision and Robotics*. Orlando, Florida, April 1992, Proc. of SPIE, vol 1708.

2. Blais F., Rioux M., Domey J.: "Optical Range Image Acquisition for the Navigation of a Mobile Robot". Proc. of *IEEE Int. Conf. on Robotics and Automation*, 1991.

3. Corion catalogue. 1992

4. Hartman K., Weber J.: "Multivision system for 3D measurement within camera vision". Proc of SPIE, *Sensor Fusion V*, Boston, Nov, 1992, pp. 475 – 487.

5. Hebert M., Kanade T., Krotkov E., Kweon I.S.: "Terrain Mapping for a Roving Planetary Explorer". Proc of *IEEE Robotics and Automation Conf.*, 1989, pp. 997 – 1002.

6. Jasiobedzki P.: "Active Image Segmentation using a Camera and a Range–finder." *Applications of Artificial Intelligence XI: Machine Vision & Robotics*. Orlando, Florida, April 1993, pp. 92 –99.

7. Jenkin M., Milios E., Jasiobedzki P., Bains N., Tran K.: "Global Navigation for ARK." Proc. of IEEE/RSJ *Intenational Conference on Intelligent Robots and Systems, IROS'93*, Yokohama, Japan, July, 1993, p. 10.

8. Nickerson SB., Jenkin M., Milios E., Down B., Jasiobedzki P., Jepson A., Terzopoulos D., Tsotsos J., Wilkes D., Bains N., Tran K.: "ARK: Autonomous Navigation of a Mobile Robot in a Known Environement." Proc. of Int. Conf. on *Intelligent Autonomous Systems (IAS–3)*, Feb 15 – 18, 1993, Pittsburgh, PA, pp.288 – 296.

9. Optech Systems Corp., Toronto, Canada, Laser Range–finder Model G–150.

10. Rioux M. : "Laser range finder based on synchronized scanners". *Appl. Opt.* 23 (21), 1984 : pp. 3844 – 3877.

11. Zhang G., Austin WJ, Wallace AM.: "Combining depth and intensity processing for scene description and object recognition." Proc of SPIE, *Sensor Fusion V*, Boston, Nov, 1992, pp. 49 – 59