The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

An Inexpensive 3D Scanner for Indoor Mobile Robots

Julian Ryde

Abstract— **A novel and inexpensive mechanism for acquiring 3D data suitable for indoor mobile robots is introduced. The scanner is currently capable of ranging up to 4m with an accuracy of 0.05m and costs less than 500USD. The scans cover the whole 360 degrees around the robot at heights from floor level to 1m. The data so obtained is particularly applicable to navigation, mapping and obstacle avoidance. The scanner consists of a webcam, visual encoder disk, angled laser line projector and determines range based on triangulation. The scan rate depends upon the frame rate of the camera and its rotation speed.**

I. Introduction

Mobile robots are increasingly being equipped with 3D ranging sensors, as researchers recognise the need for good sensory perception for many mobile robotic tasks.

Due to the ubiquity and relative low cost of cameras, stereoscopic methods for capturing depth information have arisen [1]. This approach, especially binocular stereo, is in part inspired by the manner of depth perception in humans and has been implemented intensively [2], [3], [4]. The main problem with these approaches is the need for texture information to aid the disparity calculation process. Uniformly coloured surfaces, such as those often found in indoor environments, are not imaged well with stereoscopic techniques. There have been a variety of attempts to solve this correspondence problem by projecting fixed laser light patterns into the scene under observation. This has been the case particularly in industrial applications where laser lines are used [5]. Machine vision techniques perform under almost ideal circumstances in factory environments because observations are made relatively close to the objects of interest which are moving at known speeds along a conveyor belt. Range determination by triangulation of laser lines has even been achieved outdoors under [6], [7]. Some have deployed laser stripers on robot hands [8] and others have addressed calibration procedures for laser striping systems [8], [9]. Structured light mechanisms [10] relying on projecting an entire image or pattern with a digital projector rather than a single line have recently become more popular with the advent of inexpensive digital projectors. The main advantage of this technique is the capture of range information for the entire image simultaneously. This makes imaging of moving objects possible at camera frame rates. Indeed such structured light systems have been used for real-time facial motion capture [11]. Unfortunately the large power requirements, relatively short range and small field of view restrict the deployment of these systems on mobile robots.

J. Ryde is with the Autonomous Systems Laboratory, CSIRO ICT Centre, Brisbane, Australia julian.ryde@csiro.au

One of the problems of this approach, and similar ones such as stereoscopic vision, is the non-linear dependency of accuracy on range.

For triangulation approaches that the error in range increases with the square of the range. Increasing the baseline separation reduces the error.

For the situations in which a stereoscopic system could be mounted on a medium sized robot, say $s = 1m$, and employing cameras with 640 pixel horizontal resolution and FOV of 30 degrees results in errors of 0.05m at a range of 8m. This error is the one pertaining to optimum conditions when observing a point on the centre line between the cameras. It compares poorly with a laser scanner which has a typical range error of less than 0.02m. Optimal triangulation would necessitate a separation of 4m which is unfeasible for a medium sized, indoor robot. Multiple cooperating robots however could establish correspondence points amongst themselves and thereby employ an observation separation closer to the optimum. Another system [12] that is capable of capturing range images up to 7.5m uses modulated illumination. This system looks promising however it is still relatively expensive (around 8000USD), not widely available and currently limited to indoor applications. A 2D spinning laser scanner that employs phase modulation is the subject of [13].

Finally, there are time of flight scanners which use the time taken for an infrared laser beam to travel to the object and back. These are commercially available in a variety of forms, point range finders, 2D scanners and, recently, some 3D scanning systems. Whilst these commercial 3D scanners can operate indoors and outdoors and have both good accuracy (0.02m) and range (30m), they are usually very expensive, in the region of (50,000USD for the velodyne). Due to their cost researchers have increasingly resorted to creating custom 3D scanners by modifying existing commercial 2D scanners. This is done either by rotating the scanner itself [14] or by rotating a mirror in front of the 2D scanner [15].

A less expensive alternative, described in this work, employs a laser line together with a camera and uses triangulation to determine the range. This system differs in that it returns 3D scans and does not nod the laser line. Nodding the laser line results in motion blur of the laser line in the image substantially reducing the its apparent brightness. The main problem with triangulation ranging is the rapid deterioration in accuracy with increasing range. However, although a scanner employing triangulation is somewhat less accurate than the modified 2D laser scanners mentioned above, the design innovations can improve its accuracy to within 0.05m up to a 4m range. Such accuracy is

978-1-4244-3804-4/09/$25.00 ©2009 IEEE

within acceptable limits for indoor, mobile robots. The cost of the described system consisting of laser line projector, camera and rotation mechanism depends of course upon the quality of the equipment. That used in these experiments was purchased for less than 500USD.

The paper is organised into the following sections. Section II explains both the hardware and underlying data processing algorithms that comprise the 3D scanner. Results from both simulation and real experiments are presented in Section III. The paper is concluded by Section IV which summarises the work and discusses possible future directions.

II. Scanner Design

As the accuracy of the range information obtained by triangulation declines rapidly as the range increases, it is important to maintain as large a baseline separation as possible. Human working environments tend to have good overhead clearance and so a vertical baseline separation in the order of 2m is often possible in such situations. In contrast horizontal baseline separation may be restricted to 0.5m or so. Additionally taller robots would have the advantage of being more suited to interaction with people as well as having better observational capabilities since there is generally more clutter lower down that could interfere with sensors positioned close to the floor.

A. Hardware

As is apparent from Section I, it is vital to minimise the error in the angle between the camera and the laser line and so the equipment described below is designed to achieve this goal. Generally in camera and laser line projection systems the laser plane is perpendicular to the line joining the camera and laser. This dependence of range error on separation angle means that it is important to maintain the angular difference between the camera and laser accurately. The errors associated with scanning a horizontal laser plane up and down are greater than those introduced by non-rigidity of the structure failing to maintain the fixed angle between the camera and laser. If relative motion occurs between the camera and the laser then the laser line image will move a certain distance across the camera image sensor in the time taken for the camera to capture an exposure. This results in blurring of the line with a consequent reduction in its effective brightness and an increase in its apparent width. Both of these effects reduce the accuracy of the measurement. It is therefore highly desirable for the angular velocity of the camera and the laser to be the same to eliminate motion blur.

Finally, as is demonstrated in [16], omnidirectional sensors are ideal for robots in that they enable them to map more effectively by improving the overlap between successive scans. Omnidirectional sensors would also prove valuable in planning and improving a robot's awareness of its location.

The constraints of vertical separation, fixed relative angular difference between camera and laser and a preference for omnidirectional sensors are the reasoning behind the scanner presented in Fig. 2. The scanner consists of a camera and laser line projector both mounted on a rotating platform with vertical baseline separation. An encoder disc is placed just below the camera in such a position that its circumference is visible at the bottom of the camera frame. The laser line is projected at an angle rather than horizontally so that a volume is swept as both the camera and laser line projector rotate about the vertical axis.

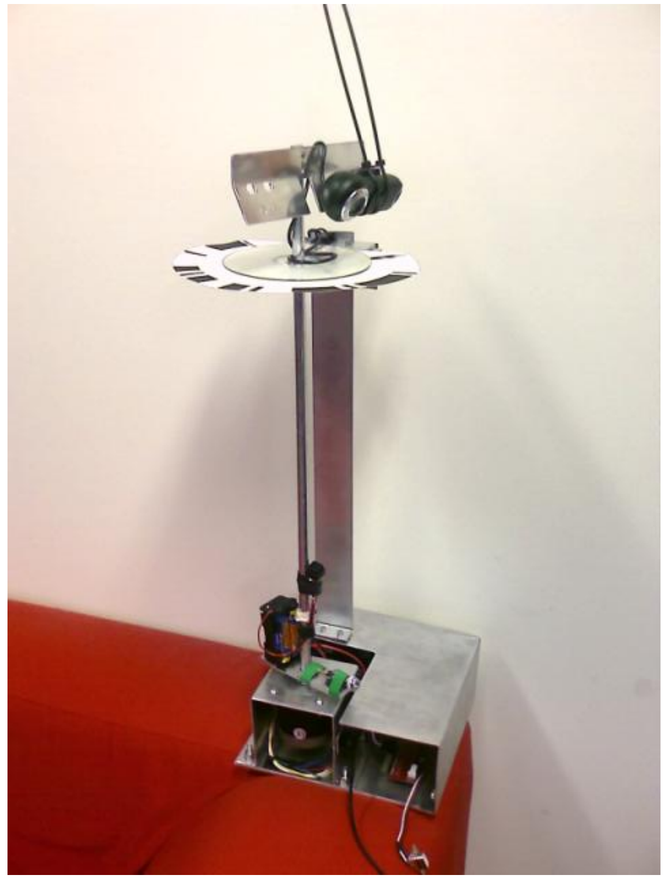

Fig. 1. Photograph of sensor consisting of a camera, visual encoder disk, rotating rod, mounting, angled laser line, stepper motor and stepper motor driver.

The two main differences of this range imaging device as compared to conventional ones based on structured light are the projection of the laser plane at an angle (as opposed to being perpendicular to the camera-laser separation vector) and the feedback of angular position via the camera image itself.

A simulated image received by the camera is shown in Fig. 2. At the bottom of the image is an encoder wheel divided into 32 sectors with binary labels. By reading the binary code and determining the offset the robot can measure the angular position of both the camera and laser to an accuracy of less than 1 degree. Indeed, if the offset is pixel accurate, the angular accuracy for the orientation of the camera approaches that of the camera's optical horizontal angular resolution.

In experiments pixel accuracy is not always achieved because the camera is focused beyond the encoder disk, about 0.1m from the lens, thus blurring its image. The question as to which point to focus the lens is important.

Since the range accuracy is much reduced at larger distances it is essential that the camera is focused out to these distances (about 4m) in order to be able to extract the angle of the laser line to the nearest pixel. If the laser line image appears blurred closer to the scanner this is not a problem because the reduction in accuracy due to blurring is more than compensated for by the increase in range accuracy due to the improved geometric configuration. For a camera with horizontal resolution of 640 pixels and field of view of 35 degrees the angular resolution of the reconstructed scene can be as good as 0.05 degrees.

The camera and laser as an entity are rotated together about the z-axis by a low cost DC motor and because rotation feedback is acquired through the image, servo or stepper motors are not needed. One scan consists of a full 360 degrees sweep around the robot always in the same direction. At the end of the scan the video capture is stopped and the scanner reset to its original position: this reciprocating motion avoids the hardware complexity introduced by continuous rotation. The camera is a standard Logitech QuickCam Pro 9000 which has good image quality due to its glass lens and an autofocus capability.

Any offset between camera and laser caused by the rotation can be reduced by increasing the rigidity of the connection between the two. And by ensuring that scans are always made in the same direction this rotational offset is consistent. For this series of experiments the camera was mounted on its side so that the wider element of the field of view was vertical thereby increasing the vertical span of the 3D scans as well as enabling the detection of objects that are close. With these conditions met each complete revolution of the camera delivered a 360 degrees horizontal 3D scan covering ranges $>$0.3m. Geometrically there is no upper limit to the range. Practically the upper limit of the range depends upon the ability of the camera to image the laser illumination of the surface, which is dependent upon the surface's reflectivity and the incident angle of the illumination.

B. Data Processing

Including the horizontal angular feedback in the image itself, rather than obtaining it by more conventional methods, confers a number of advantages. Firstly the inclusion of feedback means that the motion does not have to be so precisely controlled so there is substantial flexibility in the choice of scanning mechanisms. The main practical consideration is that the motion be sufficiently smooth to reduce the errors introduced by the rotational offset between the laser and camera, described above, to a workable minimum. This system's advantages over more conventional rotation sensors are its very high angular resolution, guaranteed data synchrony and reduced cost and complexity in the provision of rotation sensor data to the robot. Data synchrony refers to the problem of ensuring the right angular data is matched with each camera frame. Although this is simple enough in theory it can be difficult to achieve in practice, especially at higher scan speeds. The stipulation that range and angular measurements be obtained simultaneously is fulfilled by having all the necessary data included in one camera frame. Thus data collection is reduced to simply recording a video whilst scanning. The frames are extracted from the video and analysed separately. The first step in this process is to extract the pixels illuminated by the red laser line. This process is easier in darker conditions when the brightness of the laser line surpasses the ambient lighting conditions. In this situation those pixels with red (r), green(g), blue(b) components satisfying

$$r > t, g < t, b < t, \tag{1}$$

are laser illuminated pixels. The tunable threshold, t, typically has a value of around 150, on an 8 bit scale (0-255).

In bright conditions such as the image in Fig. 3 this process is more difficult and (1) no longer works. Although Fig. 3 looks relatively dark the scene was in fact brightly illuminated. This particular seen was chosen as it contains bright red regions that could generate false positives for the laser extraction process. It also clearly shows motion blur which aids the laser pixel extraction process. The motion blur affects the scene but does not affect the laser line which

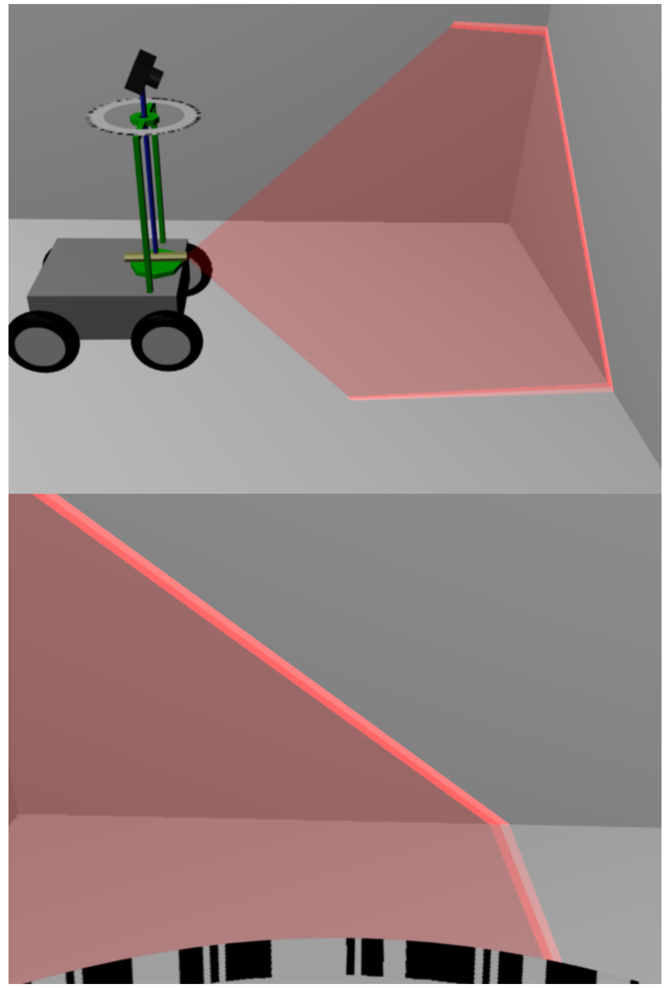

Fig. 2. Diagram of the 3D vision system operating on a robot and corresponding image from the point of view of the camera.

is rotating at the same angular velocity as the camera.

It is important to maintain colour purity so that the red component of the image responds without affecting the green and blue components. This was ensured by frame grabbing at the native resolution of the camera sensor and turning off contrast enhancement and image sharpening which many cameras perform automatically. The camera exposure was reduced to limit the laser saturating the red values and thus causing the green and blue pixels to erroneously respond to the laser light.

If the colour accuracy of the camera image is high enough the laser line should only appear in the red channel as in Fig 4. A row by row convolution with a narrow Gaussian is performed on each colour channel. Only when there is a sufficiently high response in the red channel (Fig. 5) that is absent in both the green and blue channels is the pixel designated as illuminated by the laser line. The current bearing of the camera is tracked by running edge detection across the bottom row of the image in Fig. 3. The sector can be determined from the pattern of edges corresponding to the binary code and the position within that sector is calculated from the offset. Alternatively a simpler rotionally symmetric encoder disk pattern can be used and the current sector continuously tracked during scanning. The returned pixel is then intersected with the encoder disk plane in a manner similar to that for determining the position of the 3D laser pixels, (2). The heading is calculated from the x and y components.

By projecting the laser line at an angle, different columns in the image correspond to different angular separations between the camera and laser point. This is much more accurate than actuating the laser to different angles because the relative angle between the laser and the camera is rigidly fixed.

The 3D position of each point in the image is determined by intersecting the camera ray with the laser plane. The camera ray starts at the camera position and the angle is calculated from the pixel row and column. The laser plane may be established in a number of ways but in this case it is defined by the coordinates of three points. These may be obtained by projecting the line onto a vertical wall and measuring the position of two illuminated points on the wall relative to the laser aperture. The coordinates of these two points coupled with the position of the laser aperture itself provide enough information to specify the plane. Once the laser plane is known it is then possible to calculate the intersection point with the camera ray corresponding to a laser illuminated pixel, as will now be described.

In general a plane P can be defined by a point in the plane P_X and the normal to the plane P_N. A line L can be defined in a similar manner such that L_X is a point on the line and L_D is a vector parallel to the line. Thus the point of intersection X is given by

$$X = L_X + \frac{(P_X - L_X) \cdot P_N}{L_D \cdot P_N} \cdot L_D \tag{2}$$

Each illuminated pixel with pixel coordinates (I_r, I_c) in the camera image with C_c columns and C_r rows corresponds to a camera ray with direction

$$R = \begin{bmatrix} 1 \\ C_\mu(C_c/2 - I_c) \\ C_\mu(C_r/2 - I_r) \end{bmatrix}, \tag{3}$$

where C_μ is the camera calibration constant which is calculated as

$$C_\mu = \frac{\tan(C_\theta/2)}{C_c} \tag{4}$$

with C_θ the camera's field of view in radians.

The ray vector, which is in the coordinate frame of the camera, is transformed via the standard rotation matrix about the y-axis into that of the laser plane where b is angle between the camera central pixel and the horizontal xy-plane, to give,

$$\begin{pmatrix} \cos(b) & 0 & \sin(b) \\ 0 & 1 & 0 \\ -\sin(b) & 0 & \cos(b) \end{pmatrix} R + \begin{pmatrix} 0 \\ 0 \\ s \end{pmatrix}. \tag{5}$$

Where s is the vertical separation between the camera and the laser plane. This ray is then intersected with the laser plane and the coordinates of intersection calculated by (2). The system is calbrated by scanning a vertical wall and adjusting the parameters until the wall is vertical, flat and detected at the correct distance.

III. Results

A. Simulation

Images were simulated by modelling both the scanner and robot in a 3D modelling package. The scanner model was animated to scan and images then rendered from the point of view of the camera. The advantage of a modelled environment is that the true values of the range, the position of the scanner and its various parameters are known to an arbitrary precision. These simulation experiments indicated the viability of the proposed approach as well as allowing optimisation of the mechanical set-up.

B. Real Data

An example camera image from the operating scanner is displayed in Fig. 3. The bottom of Fig. 3 shows the stationary encoder wheel. The position of the wheel is determined using edge detection and this was found to work satisfactorily despite the blurring. Also, as can be seen from Fig. 3, despite the rotation motion of the scanner the laser line image is sharp with very little motion blur. This is due to the smoothness of the surface being scanned and an absence of relative rotation between the laser and the camera.

Examples of the resultant 3D data obtained from single scans are displayed in Fig 6 and Fig. 7 in which are presented overhead views of the point clouds generated and with the points coloured by their height from the floor. From these figures it is possible to see the diagonal line stripes each of which correspond to one camera frame. The overhead views in Fig. 6 and Fig. 7 are presented to give an indication of the scanner accuracy. The walls are vertical and should appear

Fig. 4. RGB colour components of single image illustrating how the red laser line is still slightly apparent in the green component but not the blue.

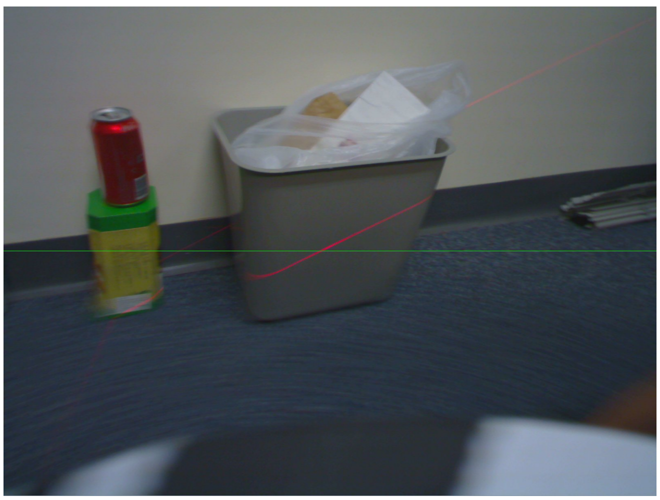

Fig. 3. Example single image, during bright conditions. Red objects present in the scene do not confuse the laser line extraction process. The green line indicates the row profiled in Fig. 5. Note how motion blur affects the scene but not the laser which co-rotating with the camera.

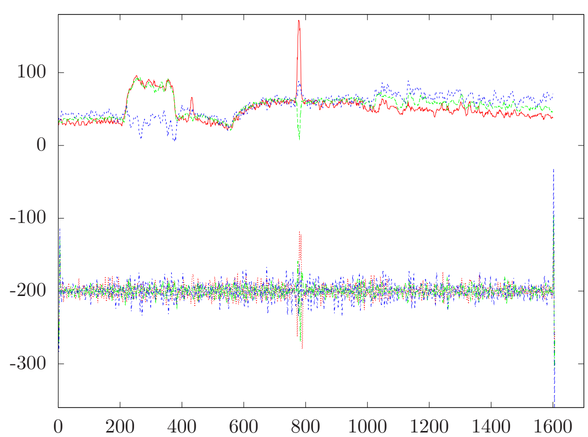

Fig. 5. Plot of the RGB components of the row highlighted in Fig. 3. Underneath the result of convolving the individual components with a Guassian. The laser line peak in the is clearly evident in the convolution of the red channel only.

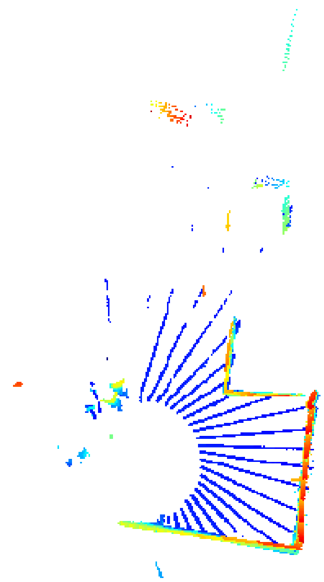

Fig. 6. Overhead view of a single scan from the camera scanner. Points are coloured by height.

as thin lines in the overhead view. The width of these lines represents the random error and it can be seen that this error is less than 0.05m. The radial lines correspond to data from the floor for each camera frame. This scan is made from approximately 50 frames and took 2 seconds to acquire with a 25 frames per second camera.

IV. Conclusions and Future Work

Described herein is an inexpensive readily constructed 3D scanner that has been designed for indoor mobile robots. The scanner returns range data from all horizontal directions and vertically from floor height to approximately 1m. The scan rate for the experiments was 2s however this can be easily configured with a trade off between scan density and time. With a camera running at 25 frames per second the scanner is capable of measuring 16,000 ranges per second. Although the accuracy deteriorates over ranges greater than 4m, the laser scanner described here offers some significant advantages.

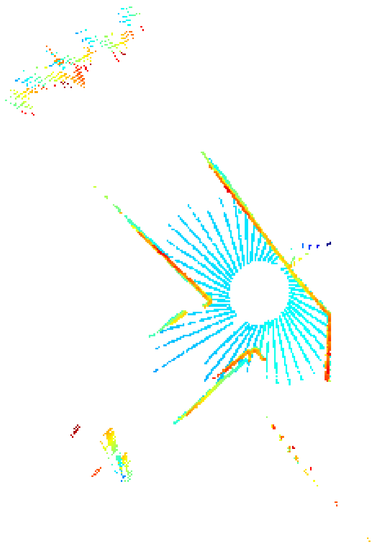

Fig. 7. Overhead view of a scan in a typical corridor (1.5m wide). The points are coloured by height with those on the floor light blue. The scan is from ground level to about 0.8m high.

First and foremost it is much less expensive than commercial time of flight scanning laser range finding systems. It also has the capability to return the texture information associated with each range reading. This colour information is useful not only to the mobile robot but also to those responsible for interpreting the data and matching it to what they themselves observe. It is accurate over the range most often required of indoor mobile robots to a level necessary to perform the usual tasks required of such robots. The camera sensor may also be exploited by standard vision algorithms, where this is necessary, for face recognition for example, or when the robot is being tele-operated. Without the laser line, panoramic images can be acquired as the multiple frames can be easily merged since the camera angle for each frame is accurately known.

Finally work is continuing on combining the colour information in each image with the ranges determined for the pixels in order to generate fully textured 3D scans and maps.

ACKNOWLEDGMENTS

The author would like to thank Nick Hillier and John Ryde for their valuable review and feedback. I would also like to express my gratitude to Rhys Davis for construction of the sensor.

REFERENCES

[1] N. Ayache and F. Lustman, "Trinocular stereo vision for robotics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 1, pp. 73–85, 1991.

[2] J. Meguro, J. Takiguchi, Y. Amano, and T. Hashizume, "3D reconstruction using multibaseline omnidirectional motion stereo based on gps/dead-reckoning compound navigation system," *Int. Journal of Robotics Research*, vol. 26, pp. 625–636, 2007.

[3] H. P. Moravec, "Robot spatial perception by stereoscopic vision and 3D evidence grids," 1996.

[4] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *IEEE Workshop on Stereo and Multi-Baseline Vision*, 2001.

[5] J. Davis and X. Chen, "A laser range scanner designed for minimum calibration complexity," *3-D Digital Imaging and Modeling, 2001. Proceedings. Third International Conference on*, pp. 91–98, 2001.

[6] K. Yoshida and S. Hirose, "Laser triangulation range finder available under direct sunlight," in *Proc. of the IEEE Int. Conf. on Robotics and Automation*, 1988, pp. 1702–1708.

[7] C. Mertz, J. Kozar, J. Miller, and C. Thorpe, "Eye-safe laser line striper for outdoor use," *IV 2002, IEEE Intelligent Vehicle Symposium*, 2002.

[8] G. Agin, "Calibration and use of a light stripe range sensor mounted to the hand of a robot," Robotics Institute, Pittsburgh, PA, Tech. Rep. CMU-RI-TR-85-20, November 1985.

[9] I. Reid, "Projective calibration of a laser-stripe range finder," *Image and Vision Computing*, vol. 14, pp. 659–666, 1996.

[10] D. Caspi, N. Kiryati, and J. Shamir, "Range imaging with adaptive color structured light," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 5, pp. 470–480, 1998.

[11] S. Zhang and P. Huang, "High-resolution, real-time 3D shape acquisition," *Computer Vision and Pattern Recognition Workshop, 2004 Conference on*, pp. 28–28, June 2004.

[12] T. Oggier, M. Lehmann, R. Kaufmann, M. Schweizer, M. Richter, P. Metzler, G. Lang, F. Lustenberger, and N. Blanc, "An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (SwissRanger)," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, ser. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, L. Mazuray, P. J. Rogers, and R. Wartmann, Eds., vol. 5249, Feb. 2004, pp. 534–545.

[13] H.Kawata, A.Ohya, S.Yuta, W.Santosh, and T.Mori, "Development of ultra-small lightweight optical range sensor system," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2005, pp. 1078–1083.

[14] A. Nüchter, K. Lingemann, J. Hertzberg, and H. Surmann, "Heuristic-based laser scan matching for outdoor 6D SLAM," in *Advances in Artificial Intelligence. 28th annual German Conf. on AI*, Sept. 2005.

[15] J. Ryde and H. Hu, "3D laser range scanner with hemispherical field of view for robot navigation," in *Proc. of IEEE/ASME Advanced Intelligent Mechatronics Int. Conf. on Mechatronics and Automation (AIM 2008)*, Xi'an, China, Jul. 2008.

[16] ——, "Mobile robot 3D perception and mapping with multi-resolution occupancy lists," in *Proc. of IEEE Int. Conf. on Mechatronics and Automation (ICMA 2007)*, Harbin, Heilongjiang, China, Aug. 2007.