Correction of Scanning and Projection Errors in An Active Depth Sensor

G. Zhang, Y. Gallaher and A. M. Wallace
Department of Computing and Electrical Engineering
Heriot-Watt University, Edinburgh EH14 4AS, Scotland

Abstract

A procedure for correction or recalibration of a laser sensor for depth image acquisition by triangulation is presented. This takes into account misalignment of both the motion of a translation stage and the geometry of the projected laser plane. The improvement in the accuracy of the acquired depth and of the registration of depth and intensity data is demonstrated.

1 Introduction

Triangulation of a projected laser spot or line with a viewing camera (or cameras) is a well established method of depth data acquisition [2]. To acquire a two dimensional depth image, a *scanning* mechanism is also required. This may be accomplished either by scanning the laser across the static object, using typically a rotating mirror or prism, or by moving the object laterally through the plane of projection of the line, using typically a translation stage. Here we present a method for calibration and error correction of both the laser geometry and the scanning motion following an initial static calibration. We consider object translation as this is the method we employ regularly [1].

2 Static calibration

We use Tsai's coplanar method [3] to establish the calibration parameters of each camera in the depth sensor whose geometry is illustrated in Figure 1. This employs two cameras to determine depth data by triangulation with a single, projected laser plane. The calibration pattern is a grid of black dots of equal diameter and spacing. The extrinsic parameters transform a point (x_w, y_w, z_w) in the world coordinate system to a point (x_c, y_c, z_c) in the camera coordinate system.

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \mathbf{R} \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + \mathbf{T} \tag{1}$$

where $\mathbf{R}$ is a rotation matrix and $\mathbf{T}$ is a translation vector

$$\mathbf{R} = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \quad \text{and} \quad \mathbf{T} = \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \tag{2}$$

BMVC 1996 doi:10.5244/C.10.69

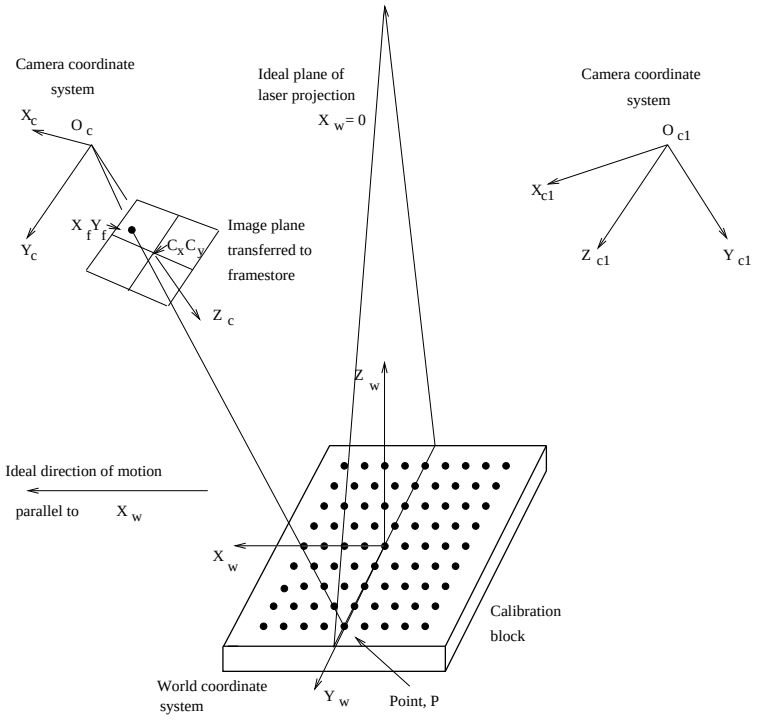

Figure 1: The sensor geometry showing world and camera coordinate systems.

The calibration procedure also provides several intrinsic parameters which allow us to determine the image plane coordinates (x_u, y_u) from the digitised image using the focal length of the lens, f, and known sensor geometry and allowing for radial distortion and sampling [3]. In principle, no further calibration is necessary provided

- The direction of the table movement is parallel to the X_w axis.
- The laser sheet is in the plane $x_w = 0$.

Assuming that the laser plane is $x_w = 0$, depth coordinates are expressed in the simplified form [4]

$$\begin{aligned} z_w &= [(r_2r_7 - r_1r_8)y_u + (r_4r_8 - r_5r_7)x_u + (r_2T_z - r_8T_x)y_u + (r_8T_y - r_5T_z)\,x_u + \\ &\quad f(r_5T_x - r_2T_y)]/A \\ y_w &= [(r_3r_7 - r_1r_9)Y_u + (r_4r_9 - r_6r_7)X_u + (r_3T_z - r_9T_x)Y_u + (r_9T_y - r_6T_z)X_u + \\ &\quad f(r_6T_x - r_3T_y)]/A \end{aligned} \tag{3}$$

where $A = (r_3r_8 - r_2r_9)y_u + (r_5r_9 - r_6r_8)x_u + f(r_2r_6 - r_3r_5)$. As the object is moved, so the x-coordinate of the depth image is defined by the stepped motion of the translation table, not by x_w which is always 0.

3 Calibration of the translation table motion

Ideally, the motion of the translation table is along the X_w axis. In practice, due to misalignment or errors in the motion of the translation table, the motion vector may be at an arbitrary direction with respect to the world coordinate system. Without loss of generality, this can be represented by a rotation of the coordinate system by α about the Z_w axis and a rotation β about the Y_w axis, A new *motion coordinate system* has the origin coincident with the world coordinate system, the X_m axis parallel to the true translation table motion, and the orthogonal Y_m and Z_m axes defined by the α and β rotations. $\mathbf{P_1}$ and $\mathbf{P_2}$ are two calibration points separated by a distance D_m which is an integral multiple of the point separation distance, L. When the calibration pattern is translated along the X_m axis by a distance D_m, the new positions of $\mathbf{P_1}$ and $\mathbf{P_2}$ are defined by $\hat{\mathbf{P}}_\mathbf{1}$ and $\hat{\mathbf{P}}_\mathbf{2}$. Ideally, $\hat{\mathbf{P}}_\mathbf{1}$ has the same coordinates as $\mathbf{P_2}$ but in practice there is an error caused by misalignment of the motion axis, as shown in Figure 2(left).

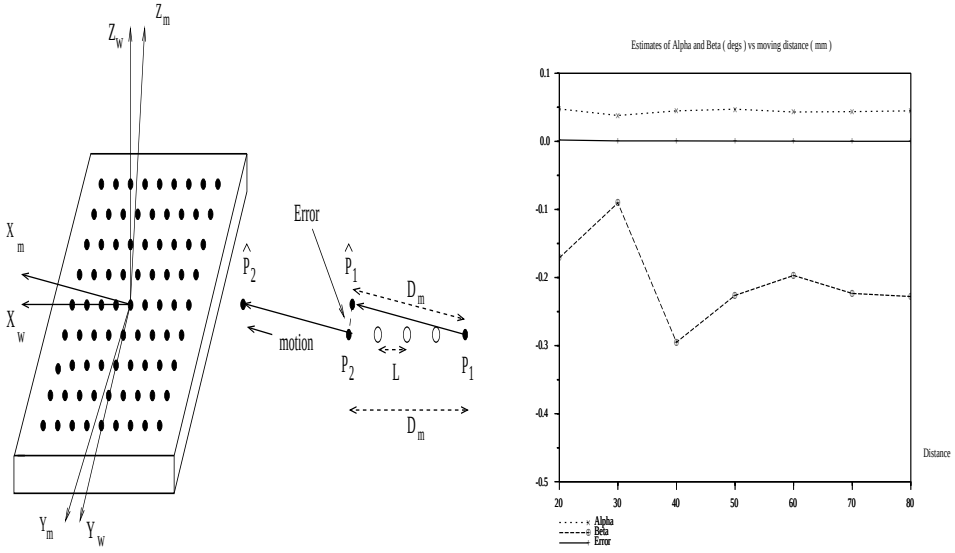

Figure 2: (left) The world and motion coordinate systems (right) Variation of α, β and the least square estimation error with motion

We define, before motion:
$(x_{w1},\ y_{w1},\ z_{w1})$, $(x_{w2},\ y_{w2},\ z_{w2})$: the world coordinates of $\mathbf{P_1}$ and $\mathbf{P_2}$.
$(x_{m1},\ y_{m1},\ z_{m1})$, $(x_{m2},\ y_{m2},\ z_{m2})$: the motion-axis coordinates of $\mathbf{P_1}$ and $\mathbf{P_2}$.
$(x_{c1},\ y_{c1},\ z_{c1})$, $(x_{c2},\ y_{c2},\ z_{c2})$: the camera coordinates of $\mathbf{P_1}$ and $\mathbf{P_2}$.
$(x_{u1},\ y_{u1})$, $(x_{u2},\ y_{u2})$: the undistorted image coordinates of $\mathbf{P_1}$ and $\mathbf{P_2}$.
and after motion:
$(\hat{x}_{w1},\ \hat{y}_{w1},\ \hat{z}_{w1})$, $(\hat{x}_{w2},\ \hat{y}_{w2},\ \hat{z}_{w2})$: the world coordinates of $\hat{\mathbf{P}}_\mathbf{1}$ and $\hat{\mathbf{P}}_\mathbf{2}$.
$(\hat{x}_{m1},\ \hat{y}_{m1},\ \hat{z}_{m1})$, $(\hat{x}_{m2},\ \hat{y}_{m2},\ \hat{z}_{m2})$: the motion-axis coordinates of $\hat{\mathbf{P}}_\mathbf{1}$ and $\hat{\mathbf{P}}_\mathbf{2}$.
$(\hat{x}_{c1},\ \hat{y}_{c1},\ \hat{z}_{c1})$, $(\hat{x}_{c2},\ \hat{y}_{c2},\ \hat{z}_{c2})$:] the camera coordinates of $\hat{\mathbf{P}}_\mathbf{1}$ and $\hat{\mathbf{P}}_\mathbf{2}$.
$(\hat{x}_{u1},\ \hat{y}_{u1})$, $(\hat{x}_{u2},\ \hat{y}_{u2})$: the undistorted image coordinates of $\hat{\mathbf{P}}_\mathbf{1}$ and $\hat{\mathbf{P}}_\mathbf{2}$.

The error between the world and motion coordinate systems is defined by the matrix $\mathbf{R_1}$, which corresponds to a rotation of α about the z-axis followed by a rotation of β about the y-axis as a linear motion only has two degrees of freedom.

$$\begin{bmatrix} x_m \\ y_m \\ z_m \end{bmatrix} = \mathbf{R_1} \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} \tag{4}$$

where

$$\mathbf{R_1} = \begin{bmatrix} \cos\alpha\cos\beta & -\sin\alpha\cos\beta & \sin\beta \\ \sin\alpha & \cos\alpha & 0 \\ -\cos\alpha\sin\beta & \sin\alpha\sin\beta & \cos\beta \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \tag{5}$$

We can now find the relative 3D displacement between $\mathbf{P_2}$ and $\hat{\mathbf{P}}_\mathbf{1}$. From equation (4) we have (before motion),

$$\begin{bmatrix} x_{m1} \\ y_{m1} \\ z_{m1} \end{bmatrix} = \mathbf{R_1} \begin{bmatrix} x_{w1} \\ y_{w1} \\ z_{w1} \end{bmatrix} \tag{6}$$

and after motion,

$$\begin{bmatrix} \hat{x}_{m1} \\ \hat{y}_{m1} \\ \hat{z}_{m1} \end{bmatrix} = \begin{bmatrix} x_{m1} + D_m \\ y_{m1} \\ z_{m1} \end{bmatrix} = \mathbf{R_1} \begin{bmatrix} \hat{x}_{w1} \\ \hat{y}_{w1} \\ \hat{z}_{w1} \end{bmatrix} \tag{7}$$

Subtracting equation (6) from (7) and rearranging the resulted equation give:

$$\mathbf{R_1^{-1}} \begin{bmatrix} D_m \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \hat{x}_{w1} - x_{w1} \\ \hat{y}_{w1} - y_{w1} \\ \hat{z}_{w1} - z_{w1} \end{bmatrix} \tag{8}$$

From equation (1),

$$\begin{bmatrix} x_{c1} \\ y_{c1} \\ z_{c1} \end{bmatrix} = \mathbf{R} \begin{bmatrix} x_{w1} \\ y_{w1} \\ z_{w1} \end{bmatrix} + \mathbf{T} \quad \text{and} \quad \begin{bmatrix} \hat{x}_{c1} \\ \hat{y}_{c1} \\ \hat{z}_{c1} \end{bmatrix} = \mathbf{R} \begin{bmatrix} \hat{x}_{w1} \\ \hat{y}_{w1} \\ \hat{z}_{w1} \end{bmatrix} + \mathbf{T} \tag{9}$$

Differencing and rearranging the above equation gives

$$\mathbf{R^{-1}} \begin{bmatrix} \hat{x}_{c1} - x_{c1} \\ \hat{y}_{c1} - y_{c1} \\ \hat{z}_{c1} - z_{c1} \end{bmatrix} = \begin{bmatrix} \hat{x}_{w1} - x_{w1} \\ \hat{y}_{w1} - y_{w1} \\ \hat{z}_{w1} - z_{w1} \end{bmatrix} \tag{10}$$

Substituting equation (10) into (8) and using the properties $\mathbf{R_1^{-1}} = \mathbf{R_1^t}$ and $\mathbf{R^{-1}} = \mathbf{R^t}$

$$\mathbf{R_1^t} \begin{bmatrix} D_m \\ 0 \\ 0 \end{bmatrix} = \mathbf{R^t} \begin{bmatrix} \hat{x}_{c1} - x_{c1} \\ \hat{y}_{c1} - y_{c1} \\ \hat{z}_{c1} - z_{c1} \end{bmatrix} \tag{11}$$

Expanding $\mathbf{R}, \mathbf{R_1}$ by their elements yields the following equations

$$\begin{aligned} r_{11} &= [r_1(\hat{x}_{c1} - x_{c1}) + r_4(\hat{y}_{c1} - y_{c1}) + r_7(\hat{z}_{c1} - z_{c1})]/D_m \\ r_{12} &= [r_2(\hat{x}_{c1} - x_{c1}) + r_5(\hat{y}_{c1} - y_{c1}) + r_8(\hat{z}_{c1} - z_{c1})]/D_m \\ r_{13} &= [r_3(\hat{x}_{c1} - x_{c1}) + r_6(\hat{y}_{c1} - y_{c1}) + r_9(\hat{z}_{c1} - z_{c1})]/D_m \end{aligned} \tag{12}$$

Substituting for r_{11}, r_{12}, r_{13} using equation (5), and expressing in terms of the measured image plane coordinates,

$$\begin{aligned} cos\alpha cos\beta &= [r_1(\hat{x}_{u1}\hat{z}_{c1} - x_{u1}z_{c1})\tfrac{1}{f} + r_4(\hat{y}_{u1}\hat{z}_{c1} - y_{u1}z_{c1})\tfrac{1}{f} + r_7(\hat{z}_{c1} - z_{c1})]/D_m \\ -sin\alpha cos\beta &= [r_2(\hat{x}_{u1}\hat{z}_{c1} - x_{u1}z_{c1})\tfrac{1}{f} + r_5(\hat{y}_{u1}\hat{z}_{c1} - y_{u1}z_{c1})\tfrac{1}{f} + r_8(\hat{z}_{c1} - z_{c1})]/D_m \\ sin\beta &= [r_3(\hat{x}_{u1}\hat{z}_{c1} - x_{u1}z_{c1})\tfrac{1}{f} + r_6(\hat{y}_{u1}\hat{z}_{c1} - y_{u1}z_{c1})\tfrac{1}{f} + r_9(\hat{z}_{c1} - z_{c1})]/D_m \end{aligned} \tag{13}$$

These three nonlinear equations are determined with three unknowns, α, β and $\hat{z}_{c1}$, since f and the matrix $\mathbf{R}$ are known from the static calibration, D_m is known precisely from the translation table motion, $(\hat{x}_{u1}, \hat{y}_{u1})$ and (x_{u1}, y_{u1}) are measured image plane coordinates obtained using the intrinsic parameters of the static calibration, and z_{c1} is also known since the initial measurement is taken from the calibrated position.

Equation (13) can be represented

$$\begin{aligned} cos\alpha cos\beta &= A\hat{z}_{c1} - B \\ sin\alpha cos\beta &= -C\hat{z}_{c1} + D \\ sin\beta &= E\hat{z}_{c1} - F \end{aligned} \tag{14}$$

where

$$\begin{aligned} A &= \tfrac{1}{fD_m}(r_1\hat{x}_{u1} + r_4\hat{y}_{u1} + r_7 f) & B &= \tfrac{1}{fD_m}(r_1 x_{u1}\hat{z}_{c1} + r_4 y_{u1}\hat{z}_{c1} + r_7 z_{c1} f) \\ C &= \tfrac{1}{fD_m}(r_2\hat{x}_{u1} + r_5\hat{y}_{u1} + r_8 f) & D &= \tfrac{1}{fD_m}(r_2 x_{u1}\hat{z}_{c1} + r_5 y_{u1}\hat{z}_{c1} + r_8 z_{c1} f) \\ E &= \tfrac{1}{fD_m}(r_3\hat{x}_{u1} + r_6\hat{y}_{u1} + r_9 f) & F &= \tfrac{1}{fD_m}(r_3 x_{u1}\hat{z}_{c1} + r_6 y_{u1}\hat{z}_{c1} + r_9 z_{c1} f) \end{aligned}$$

$\hat{z}_{c1}$ can be obtained by solving the quadratic equation arising from the constraint $r_{11}^2 + r_{12}^2 + r_{13}^2 = 1$, i.e. $(A\hat{z}_{c1} - B)^2 + (C\hat{z}_{c1} - D)^2 + (E\hat{z}_{c1} - F)^2 = 1$. Using several pairs of points, least square solutions of α and β were obtained by minimising the residuals of equation (13) using the Newton-Raphson method.

Following calibration, when depth data is acquired, the object is moved along the motion axis. From equations (1) and (4), the rotation matrix is replaced by $\mathbf{RR_1^{-1}}$ so that the transformation between the motion and camera coordinate systems is

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \mathbf{RR_1^{-1}} \begin{bmatrix} x_m \\ y_m \\ z_m \end{bmatrix} + \mathbf{T} \tag{15}$$

The x_m values are given by the table translation and the y_m and z_m values are calculated from the measured image coordinates using the corrected calibration parameters.

4 Calibration of the Laser Light Plane

Previously, we assumed the laser sheet was in the $x_w = 0$ plane to compute the depth coordinates from equation (3). To obtain the true orientation of the laser

sheet, we used a set of accurately machined metal plates placed in the path of the laser stripe, and acquired an image at each height. The equation of the laser plane in explicit form is given

$$x_m = a_l y_m + b_l z_m + c_l \tag{16}$$

At a known step height, z_{w0}, this becomes $x_m = a_l y_m + (b_l z_{w0} + c_l)$ which defines a line resulting from the intersection of the laser light plane and the plane $z = z_{w0}$. By changing the z_{w0} value, many lines of intersection can be obtained. Assuming m lines of intersection are acquired, p points are acquired at equally spaced positions along each line, then this can be expressed as a matrix equation ($n = mp$)

$$\mathbf{V} = \mathbf{U}\mathbf{X} \tag{17}$$

where $\mathbf{X}$ is the unknown laser sheet parameter vector and

$$\mathbf{V} = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} y_1 & z_1 & 1 \\ y_2 & z_2 & 1 \\ & \cdots & \\ y_n & z_n & 1 \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} a_l \\ b_l \\ c_l \end{bmatrix} \tag{18}$$

Equation (17) is over-determined and can be solved by least-squares regression to minimise $\|\mathbf{U}\mathbf{X} - \mathbf{V}\|$ which gives $\widehat{\mathbf{X}} = (\mathbf{U}^\mathbf{T}\mathbf{U})^{-1}\mathbf{U}^\mathbf{T}\mathbf{V}$. The full transformation equation has the following form

$$z_m = a_1 x_m + b_1 \tag{19}$$

$$y_m = a_2 x_m + b_2 \tag{20}$$

where a_1, b_1, a_2, b_2 are constants defined by the measured image plane coordinates and the static calibration. Replacing x_m in equation (19) by equation (16) and solving this equation produce the rectified world coordinate values which satisfy the laser plane constraints:

$$\begin{aligned} z_m &= [b_1 + c_l a_1 + a_l(a_1 b_2 - a_2 b_1)]/(1 - a_l a_2 - a_1 b_l) \\ y_m &= [b_2 + c_l a_2 - b_l(a_1 b_2 - a_2 b_1)]/(1 - a_l a_2 - a_1 b_l) \\ x_m &= a_l y_m + b_l z_m + c_l \end{aligned} \tag{21}$$

When $a_l = b_l = 0.0$, this is equivalent to the laser sheet parallel located at the $x_m = 0$ plane.

5 Experiments

In the triangulation range sensor illustrated schematically in Figure 1 the cameras have CCD cell sizes of 8.6μm by 8.3μm along rows and columns respectively. The effective usage of each sensor is 752 columns by 512 rows. The framegrabber resamples each row and produces images of 768 columns by 512 rows. The accuracy of the translation table movement was measured experimentally as $0.020\,mm$ for a movement of $10\,mm$, which is in agreement with the manufacturer's data.

5.1 Calibration of translation table motion

First, each camera was calibrated. Then, the calibration pattern was moved by a known distance to compute motion direction. In principle, an estimate of the motion error matrix can be obtained from the motion of a single point. In practice, we used all the points (from 81 to 118 pairs of points) and then moved the translation table by a displacement D_m. The obtained images were then processed to extract the image coordinate points to subpixel accuracy. Results from the right camera with D_m =20.00 mm to 80.00mm are shown in Fig. 2(b). The results show that the misalignment of the translation stage is mainly around the Y_m axis, β, which is approximately -0.22 degrees for large movements. The misalignment around the Z_m axis is 0.04 degrees.

To verify the correctness of the values obtained for α and β, the rotation angles were applied to rectify the static calibration parameters. We computed the movement distance of the all the points used for obtaining the motion orientation in the motion coordinate system. Since it is known that all these points lie in the plane $z_w = 0$, the x_w, y_w and thereafter (x_m, y_m, z_m) values can be calculated using equations (4). Fig. 4 (left) shows the movement distances computed using the corrected and the assumed motion orientation information, compared to the actual movement distance. It is seen that the use of new calibration parameters has provided clear improvement to the accuracy of the positional values x_m. As these define the x-coordinates of the depth image, this is important.

5.2 Calibration of laser sheet geometry

The laser plane was aligned manually with respect to the calibration pattern. Five positions of the calibration plates were used to perform the laser plane calibration described in section 4. The fitted coefficients for the laser plane were

a_l	b_l	c_l
3.556e-3	-1.114e-1	-4.097e-1

The calculated angles between the laser plane and the Y_m and Z_m axes were 0.20 and 6.36 degrees respectively, which illustrates the extent of inaccuracy in the initial alignment. To show the effectiveness of the procedure, we extracted both depth and the x_m values from laser stripes impinging on a series of calibration plates of known height.

The results illustrated in Figure 3 show that the use of the laser post-calibration procedure provides a correction of the systematic error in the depth data, despite the random fluctuations in the raw data due to the unevenness of the laser width and intensity along the stripe. The difference in y_m is insignificant with the maximum difference within the chosen range less than $0.1mm$. However, there is large disparity in the z_m measurements as shown in Figure 3. This improvement can be as large as $11mm$ depending on the stand-off distance.

Three further sets of measurements were obtained from a single laser stripe on a calibration plate at a fixed depth. The first used corrected motion orientation and the corrected laser sheet plane. The second used the assumed motion and assumed laser sheet plane. The third used the assumed motion but the corrected laser sheet

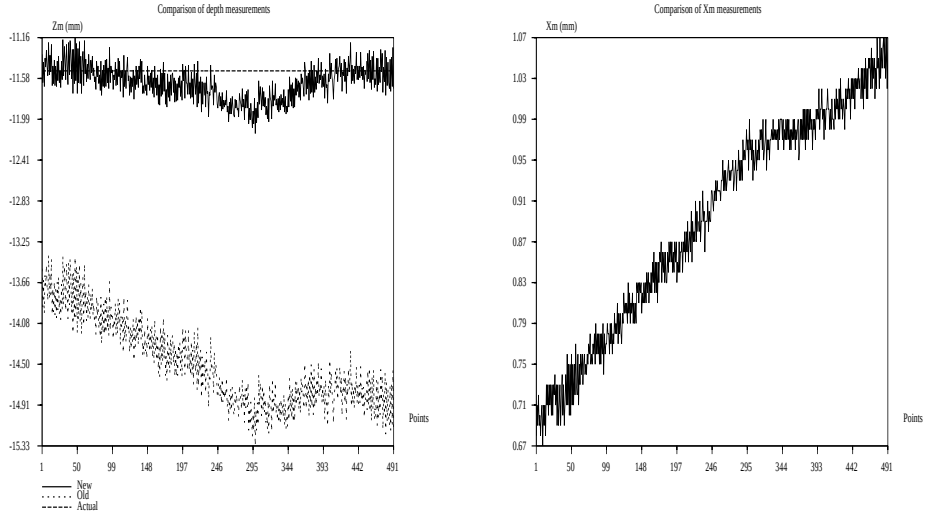

Figure 3: Depth and x_m values measured along a laser stripe at $z_m = -11.50mm$. On the left, the solid curve shows the depth computed with motion and laser plane correction. The dotted curve shows the depth measurements without motion and laser plane correction. The dashed line shows the actual depth value. On the right, the solid curve shows the x_m values computed along a stripe with motion and laser plane correction whereas $x_m = 0$ is assumed without motion and laser plane correction.

plane. The differences between the median of these depth measurements and the actual depth are illustrated in Figure 4 (right) As expected, this shows that the largest source of error is due to misalignment of the incorrect laser plane which proves more difficult in practice. The closer a point is to the camera, the larger the error. This error can be as large as $11mm$ at the height of $33.84mm$ using the manually aligned laser sheet.

5.3 An example of registration

A comparison of the improvement to depth data quality with motion direction and laser plane correction is illustrated in Figure 5. An object with two planes forming a roof edge was chosen, and two dot patterns pasted on these two planes. Both the sizes of these dots and the spacing are $5mm$. The two planes have varying depth ranges of $55mm$ and the distance between cameras and the base plane was about $600mm$. The initial striping position was chosen so that it touched the edges of those black dots in a column. The left image shows good registration along both X_m and Y_m dimensions though a slight mismatch on the right part of the image is noticeable due to the accumulated table movement error. The right image shows much poorer registration. This happens mainly along the X_m dimension which is consistent with our analysis that any error in the assumed laser plane has little

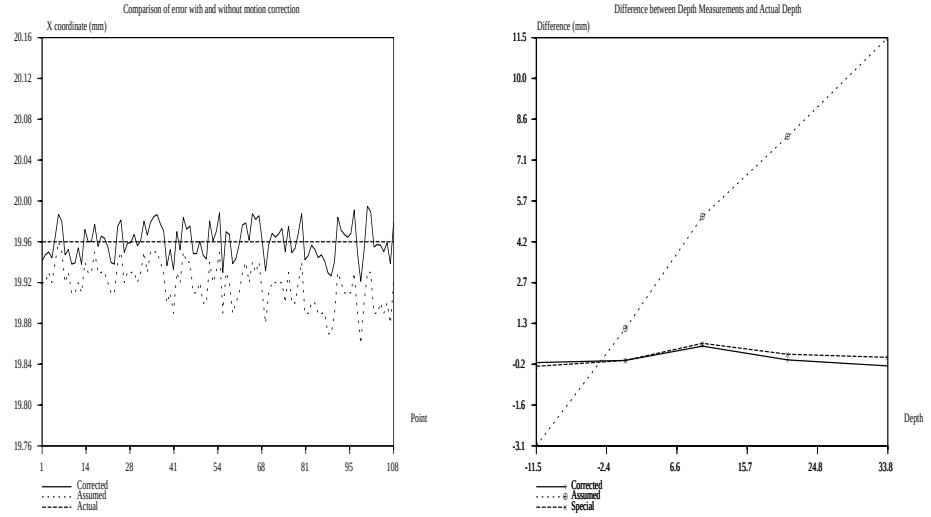

Figure 4: (left) Motion distances computed using rectified and the assumed motion orientation information are represented as solid and dotted lines respectively. The dashed line shows the actual moving distance. Moving distance is $20.00mm$. (right) The difference between the median of depth measurements along a single stripe and the ground truth using different parameters: with motion and laser sheet correction (solid line); without motion and laser sheet correction (dotted line); without motion correction, but with laser sheet correction (dashed line).

effect on the calculation of the y_m coordinate.

6 Discussion

We have shown how errors in alignment of the scanned motion of an object and in the projection of a laser plane can cause systematic errors in acquired depth images from triangulation sensors. To allow for these factors, we have introduced a second stage of calibration or error correction, to find the difference between the assumed and actual parameters of the scanning motion and the laser plane. Applying these corrections, we have been able to acquire depth measurements which are much closer to the true values, demonstrated by the use of calibration plates of known thickness, and by registration of corrected depth and intensity images in the xy plane of an intensity camera.

Acknowledgements

The work is supported by UK EPSRC Grant GR/J07891: Part Identification and Positioning by Sensor Fusion.

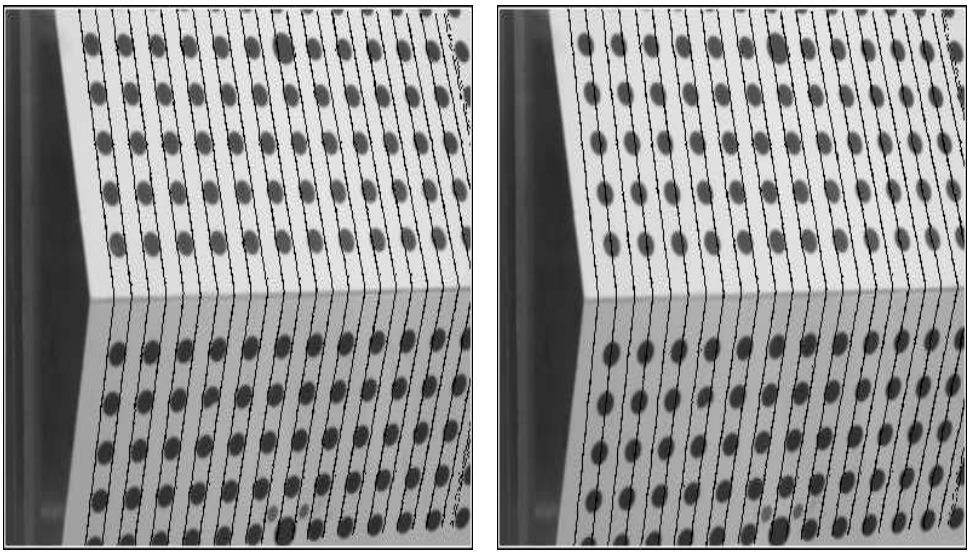

Figure 5: An example of registered depth and intensity images, with laser stripes (black lines) superimposed onto the intensity image to show the registration, (left) with motion direction and laser plane correction, and (right)with assumed motion direction and laser plane. Ideally, the laser lines are tangential to the circumference of the circles.

References

[1] J Clark, G Zhang, and A.M. Wallace. Image acquisition using fixed and variable triangulation. In *Proc. of IEE Int. conf on Image Processing*, pages 539–543, 1995.

[2] D. Nitzan. Three-dimensional vision structure for robot applications. *IEEE Trans PAMI*, 10(3):291–309, 1988.

[3] R. Y. Tsai. A versatile camera calibration technique for high accuracy 3d machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Trans. on Robotics and Automation*, 3(4):323–344, 1987.

[4] R.G. Willson. Modeling and calibration of automated xoom lenses. In *Ph.D Thesis*. Carnegie-Mellon University, 1994.