# OPTIMIZING THE EXCAVATOR WORK BY SENSORS

**Phd.Constantin P. Bogdan,** University of Petrosani, ROMANIA
**Prof.Ph.Nan Marin Silviu,**University of Petrosani, ROMANIA

**Abstract:** Recovering the three dimensional structure of a scene accurately and robustly is important for object modelling and robotic grasp planning, which in turn are essential prerequisites for grasping unknown objects in a cluttered environment. Shape recovery techniques are broadly described as either passive or active. Passive methods include recovering shape from a single image using cues such as shading, texture or focus, and shape from multiple views using stereopsis or structure-from-motion. Passive shape recovery has relatively low power requirements, is non-destructive and more akin to our biological sensing modalities. However, the accuracy and reliability of passive techniques is critically dependent on the presence of sufficient image fea tures and the absence of distractions such as reflections.

**Keywords:** Boole algebras

Mathematics Subject Classification 2000: 54A40

Active sensing relies on projecting energy into the environment and interpreting the modified sensory view. Active techniques generally achieve high accuracy and reliability, since the received signal is well-defined but at the cost of high power consumption. Active shape recovery is generally based on either some form of structured light projection or time-of-flight ranging. Recent developments in active CMOS time-of-flight cameras enable a complete 3D image to be captured in a single measurement cycle however this technology is still limited by timing resolu- tion. Structured light ranging is based on the same principles as stereopsis, but with a masked light pattern projected onto the scene to provide easily identifiable image features. The main research in this area has focused on designing coded light patterns to uniquely identify each surface point. For robotic applications, the disadvantage of this approach is the requirement of a large, high-power pattern projector.

Light stripe ranging is a simple structured light technique that uses a light plane (typically generated using a laser diode) to reconstruct a single range profile of the target with each frame. This technique offers the advantages of compactness, low power consumption and minimal computational expense. However, applying light stripe sensing to service robots presents unique challenges. Conventional scanners require the brightness of the stripe to exceed that of other features in the image to be reliably detected, making the technique most effective in highly controlled environments. Robust light stripe detection methods have been proposed in previous work but suffer from issues including assumed scene structure, lack of error recovery and acquisition delay. Thus, the goal in this chapter is to develop a robust light stripe sensor for service robots that overcomes this major shortcoming.

The solution is an actively calibrated stereoscopic light stripe scanner, capable of robustly detecting the stripe in the presence of

secondary reflections, cross-talk and other noise mechanisms. Measurement validation and noise rejection are achieved by exploiting the redundancy in stereo measurements (conventional scanners use a single camera). The validation and reconstruction algorithms are optimal with re spect to sensor noise, which results in higher precision ranging than conventional methods. Furthermore, selfcalibration from an arbitrary non-planar target allows ro- bust validation to be achieved independently of ranging accuracy.

The following section introduces the basic principles of conventional single camera light stripe range sensing and discusses existing robust techniques .

## Main Results

### 1 Conventional Light Stripe Ranging and Related Work

Light stripe ranging is an active, triangulation-based technique for non-contact surface measurement that has been studied for several decades . A review of conventional light stripe scanning and related range sensing methods . Range sensing is an important component of many robotic applications, and light stripe ranging has been applied to a variety of robotic tasks including navigation , obstacle detection , object recognition for grasping and visual servoing.

Figure 1 illustrates the operation of a conventional single-camera light stripe sensor. The principle is similar to binocular stereo with one of the cameras replaced by a light plane projector (typically a laser with a suitable lens). The stripe reflected from the target is measured by the camera and each point in the 3D surface profile (for example **X** in Figure 1) is reconstructed by triangula- tion, using the known transformation between the camera and projector. To capture a complete range image, the light plane is mechanically panned across the target and the range slices are registered into a mesh.

The drawback of conventional single-camera light stripe ranging is that favourable lighting conditions and surface reflectance properties are required so the stripe can be identified as the brightest feature in the image. In practice, this is achieved by coating the target with a matte finish, using high contrast cameras or reducing the level of ambient light. When the range sensor is intended for use by a service robot to recognize unknown objects in a domestic or office environment , various noise mechanisms may interfere to defeat stripe detection: smooth surfaces cause secondary reflections, edges and textures may have a stripe-like appearance, and cross-talk can arise when multiple robots scan the same environment.

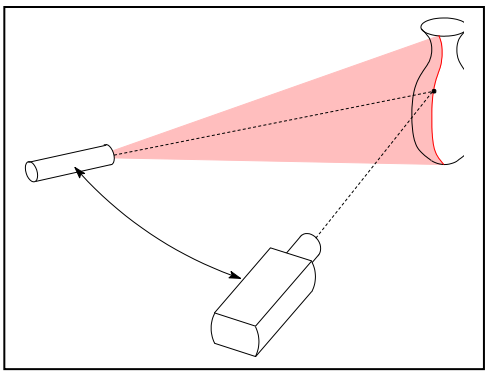

Fig. 1. Conventional single-camera light stripe sensor

A number of techniques for improving the robustness of light stripe scanners have been proposed in other work, using both stereo and single-camera configurations. Magee et develop a scanner for industrial inspection using stereo measurements of a single stripe. Spurious reflections are eliminated by combining stereo fields via a minimum intensity operation. This technique depends heavily on user intervention and a priori knowledge of the scanned target. Trucco et al. Use stereo cameras to measure a laser stripe, and treat the system as two in- dependent single-camera sensors. Robustness is achieved by imposing a number of consistency checks to validate the range data, the most significant of which requires independent single-camera reconstructions to agree within a threshold distance. An- other constraint requires valid scan-lines to contain only a single stripe candidate, but a method for error recovery in the case of multiple candidates is not proposed. Thus, secondary reflections cause both the true

and noisy measurements to be rejected.

Nakano et develop a similar method to reject false data by requiring consensus between independent scanners, but using two laser stripes and only a single camera. In addition to robust ranging, this configuration provides direct measure- ment of the surface normal. The disadvantage of this approach is that each image only recovers a single range point at the intersection of the two stripes, resulting in a significant acquisition delay for the complete image.

Other robust scanning techniques have been proposed using single-camera, sing-

le-stripe configurations. Nygards and Wernersson identify specular reflections by moving the scanner relative to the scene and analyzing the motion of recon- structed range data. Periodic intensity modulation distinguishes the stripe from random noise. Both of these methods require data to be associated between multiple images, which is prone to error. Furthermore, intensity modulation does not disambiguate secondary reflections, which vary in unison with the true stripe. Alter natively, Clark et al. Use linearly polarized light to reject secondary reflections from metallic surfaces, based on the observation that polarized light changes phase with each specular reflection. However, the complicated acquisition process requires multiple measurements through different polarizing filters.

Unlike the above robust techniques, the method described below uniformly rejects interference due to secondary reflections, cross-talk, background features and other noise mechanisms. A significant improvement over previous techniques for error detection is a mechanism for the recovery of valid measurements from a set of noisy candidates. The reconstructed depth data is optimal with respect to sensor noise, unlike the stereo techniques , and stereo measurements are fused with the light plane parameters to provide greater precision than a single-camera configuration. Finally, conventional techniques using a special camera for stripe detec- tion typically require a second camera to measure colour. The ability to operate in ambient indoor light allows our sensor to measure implicitly registered colour and range in the same camera.

## 2 Robust Stereoscopic Light Stripe Sensing

As discussed above, the shortcomings of light stripe sensors arise from the difficulty in disambiguating the primary reflection of the stripe from secondary reflections, cross-talk and other sources of noise. The following sections detail the principles of an optimal strategy to resolve this ambiguity and robustly identify the true stripe by exploiting the redundancy in stereo measurements.

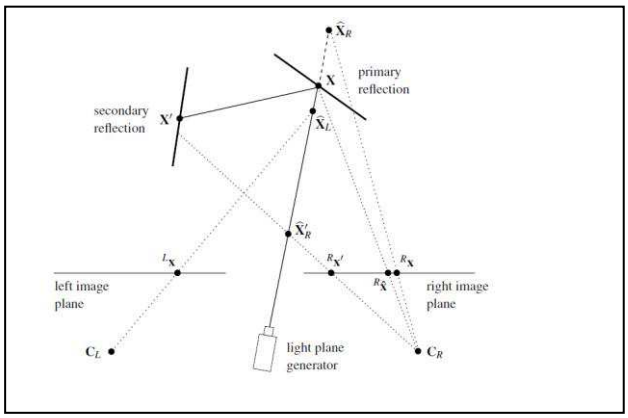


Fig.2. Validation/reconstruction problem.

A primary reflection at $\mathbf{X}$ that is measured (using a noisy process) at ${}^{L}\mathbf{x}$ and ${}^{R}\mathbf{x}$ on the stereo image planes. However, a secondary specular reflection causes another stripe to appear at $\mathbf{X}$ , which is measured on the right image plane at ${}^{R}\mathbf{x}$ but obscured from the left camera (in practice, such noisy measurements are produced by a variety of mechanisms other than secondary reflections). The 3D reconstructions, labelled $\mathbf{X}_L$, $\mathbf{X}_R$ and $\mathbf{X}_R$ in Figure 2, are recovered as the intersection of the light plane and the rays back-projected through the image plane measurements. These points will be referred to as the single-camera reconstructions. As a result of noise on the CCD (exaggerated in this example), the back-projected rays do not intersect the physical reflections at $\mathbf{X}$ and $\mathbf{X}$ .

The robust scanning problem may now be stated as follows: given the laser plane position and the measurements ${}^{L}\mathbf{x}$, ${}^{R}\mathbf{x}$ and ${}^{R}\mathbf{x}$ , one of the left/right candidate pairs, $({}^{L}\mathbf{x},{}^{R}\mathbf{x})$ or $({}^{L}\mathbf{x},{}^{R}\mathbf{x})$, must be chosen as representing stereo measurements of the pri- mary reflection. Alternatively, all candidates may be rejected. This task is referred to as the validation problem, and a successful solution in this example should identify $({}^{L}\mathbf{x},{}^{R}\mathbf{x})$ as the valid measurements. The measurements should then be combined to estimate the position of the ideal projection ${}^{R}\hat{\mathbf{x}}$ (arbitrarily chosen to be on the right image plane) of the actual point $\mathbf{X}$ on the surface of the target.

Formulation of optimal validation/reconstruction algorithms should take account of measurement noise, which is not correctly modelled in previous related work. Laser stripe measurements are validated by applying a *fi*xed thresh- old to the difference between corresponding single-camera reconstructions $(\mathbf{X}_L, \mathbf{X}_R$

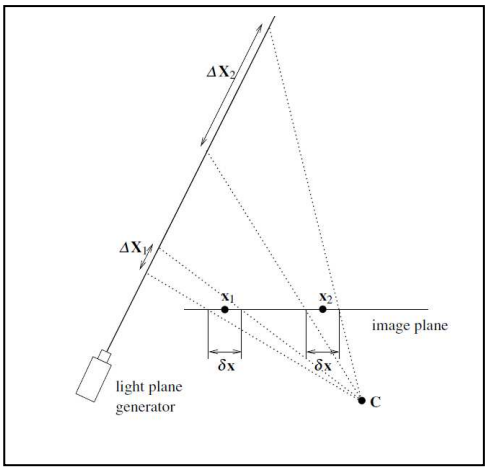


Fig.3. Variation of reconstruction error with depth.

and $\mathbf{X}_R$ in Figure 3). Such a comparison requires a uniform reconstruction error over all depths, which Figure 3.3 illustrates is clearly not the case. Two independent measurements at $\mathbf{x}_1$ and $\mathbf{x}_2$ generally exhibit a constant error variance on the image plane, as indicated by the interval $\delta\,\mathbf{x}$. However, projecting $\delta\,\mathbf{x}$ onto the laser plane reveals that the reconstruction error increases with depth, since $\Delta\mathbf{X}_1 < \Delta\mathbf{X}_2$ in Figure 3. Thus, the validation threshold on depth difference should increase with depth to account for measurement noise, otherwise validation is more lenient for closer re- constructions. Similarly, taking either $\mathbf{X}_L$ , $\mathbf{X}_R$ or the arithmetic average $^{1}(\mathbf{X}_L + \mathbf{X}_R)$ as the final reconstruction in Figure 2 is generally sub-optimal for noisy measurements.

The following sections present optimal solutions to the validation/reconstruction problem, based on an error model with the following features (the assumptions of the error model are corroborated with experimental results :

1. Light stripe measurement errors are independent and Gaussian distributed with uniform variance over the entire image plane.
2. The dominant error in the light plane position is the angular error about the axis of rotation as the plane is scanned across the target, which couples all measurements in a given image.

3. All other parameters of the sensor (described in the following section), are assumed to be known with sufficient accuracy that any uncertainty can be ignored for the purpose of validation and reconstruction.

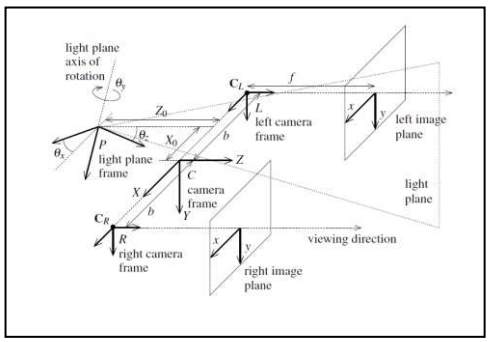


Fig. 4. Light stripe camera system model

## 3 System Model

Figure 3.4 details the parameters of the system model for the stereoscopic light stripe sensor. L and R denote the left and right camera frames, P is the frame of the rotating light plane and C is the camera frame. The cameras are modelled by the $3 \times 4$ projection matrices ${}^{L,R}P$, given by equation . The cameras are assumed to have identical focal length $f$, and are in rectilinear stereo configuration (optical axes aligned to the world z-axis) with optical centres located at $\mathbf{C}_{L,R} = (\ \ b, 0, 0)$ in the camera frame. Tthe cameras are allowed to verge about the y-axis, and projective rectification is applied to every frame to recover the equivalent rectilinear stereo measurements.

Frame P is rigidly attached to the laser, and points $\mathbf{X}$ on the light plane are defined by the plane equation $\boldsymbol{\Omega}\ \mathbf{X} = 0$, where $\boldsymbol{\Omega}$ represents the parameters of the light plane. Furthermore, frame P is defined such that the light plane is approximately vertical and parallel to the z-axis. This allows $\boldsymbol{\Omega}$ to be expressed in P as

$$^{P}\boldsymbol{\Omega} = (1,\ B_0,\ 0,\ D_0)$$

where $B_0$ is related to the small angle between the plane and the y-axis $B_0$ for an approximately vertical plane), and $D_0$ is the distance of the plane from the origin.

Thus, the normal vector of the plane is already normalized, since $|(1, B_0, 0)| \approx 1$. During a scan, frame P rotates about its y-axis with angle $\theta_Y$, where $\theta_Y \equiv 0$ when the light plane is parallel to the optical axes of the cameras. The rotation axis intersects the xz-plane of the camera frame at $(X_0, 0, Z_0)$ (in the camera frame), and the orien- tation of the rotation axis relative to the y-axis of the camera frame is defined by the small fixed angles $\theta_X$ and $\theta_Z$. The scan angle ($\theta_Y$ in Figure 4) is assumed to have a linear relationship with the measured optical encoder value e via two additional parameters m and k:

$$\theta_Y = me + k$$

Now let ${}^{C}H_P$ represent the homogeneous coordinate transformation from P to C. Representing the homogeneous transformation for rotation about the x, y and z-axes by angle $\theta$ by matrices $R_X(\theta)$, $R_Y(\theta)$ and $R_Z(\theta)$, and the translation of $(X, Y, Z)$ by the matrix $T(X, Y, Z)$, ${}^{P}H_C$ can be expressed as ${}^{C}H_P = T(X_0, 0, Z_0) \cdot R_Z(\theta_Z) \cdot R_X(\theta_X) \cdot R_Y(\theta_Y)$

It is straightforward to show that if ${}^{P}H_C$ is the coordinate transformation from P to C, the plane parameters transform from P to C as

$$^{C}\boldsymbol{\Omega} = (^{C}H_P)^{-} \cdot {}^{P}\boldsymbol{\Omega}$$

Combining equations the laser plane parameters are expressed in the camera frame as:

$$^{C}\Omega = \begin{pmatrix} c_y c_z - s_x s_y s_z - B_0 c_x s_z \\ c_y s_z + s_x s_y c_z + B_0 c_x c_z \\ -c_x s_y + B_0 s_x \\ (s_x s_y s_z - c_y c_z + B_0 c_x s_z) X_0 + (c_x s_y - B_0 s_x) Z_0 + D_0 \end{pmatrix}$$

(1)

where $c_X = \cos\theta_X$ and $s_X = \sin\theta_X$. By making the simplifying assumptions $B_0, \theta_X, \theta_Z \ll 1$, many insignificant terms in equation can be neglected to give an approximate model:

$$^{C}\Omega = \begin{pmatrix} \cos\theta_y \\ \theta_x \sin\theta_y + \theta_z \cos\theta_y + B_0 \\ -\sin\theta_y \\ -X_0 \cos\theta_y + Z_0 \sin\theta_y + D_0 \end{pmatrix}$$

(2)

Finally, equations allow points on the laser plane to be identified in the world frame, using the plane equation ${}^{C}\boldsymbol{\Omega}\ (B_0, D_0, X_0, Z_0, \theta_X, \theta_Y, \theta_Z)^{C}\mathbf{X} = 0$. exactly. The optimal reconstruction $\mathbf{X}$ projects to the ideal corresponding points on the left and right image planes at ${}^{L,R}\hat{\mathbf{x}} = {}^{L,R}P\mathbf{X}$, according to equation . Now, in a similar manner to equation , the sum of squared errors between the ideal and measured points can be used to determine whether the candidate measurement pair $(\mathbf{x}_L, \mathbf{x}_R)$ corresponds to a point $\mathbf{X}$ on the light plane:

$$E = d^2({}^{L}\mathbf{x}, {}^{L}\hat{\mathbf{x}}) + d^2({}^{R}\mathbf{x}, {}^{R}\hat{\mathbf{x}}) = d^2({}^{L}\mathbf{x}, {}^{L}P\mathbf{X}) + d^2({}^{R}\mathbf{x}, {}^{R}P\mathbf{X})$$

where $d(\mathbf{x}_1, \mathbf{x}_2)$ is the Euclidean distance between $\mathbf{x}_1$ and $\mathbf{x}_2$. For a given candidate pair, the optimal reconstruction $\mathbf{X}$ with respect to image plane error is found by a constrained minimization of E with respect to the condition that $\mathbf{X}$ is on the laser plane:

$$\boldsymbol{\Omega}\ \mathbf{X} = 0$$

When multiple ambiguous correspondences exist, equation is optimized with respect to the constraint in for all possible candidate pairs, and the pair with minimum error is chosen as the most likely correspondence. Finally, the result is validated by imposing a threshold on the maximum allowed squared image plane error E.

Performing the constrained optimization of equations is analytically cumbersome. Fortunately, the problem may be reduced to an unconstrained optimization by determining the direct relationship between projections ${}^{L}\hat{\mathbf{x}}$ and ${}^{R}\hat{\mathbf{x}}$ for points on the light plane. Taking the intersection between the light plane and the backprojected ray from ${}^{R}\hat{\mathbf{x}}$, the relationship between $\mathbf{X}$ and ${}^{R}\hat{\mathbf{x}}$ for points on the light plane $\boldsymbol{\Omega}$ is given by (see Appendix B for a complete derivation of this result):

$$\mathbf{X} = [\mathbf{C}_R({}^{R}P^{+}\,{}^{R}\hat{\mathbf{x}}) \quad -({}^{R}P^{+}\,{}^{R}\hat{\mathbf{x}})\,\mathbf{C}_R]\boldsymbol{\Omega}$$

where ${}^{R}P^{+}$ is the pseudo-inverse of ${}^{R}P$. Now, projecting $\mathbf{X}$ onto the left image plane, the relationship between the projections ${}^{L}\hat{\mathbf{x}}$ and ${}^{R}\hat{\mathbf{x}}$ for points on the light plane $\boldsymbol{\Omega}$ can be expressed as :

$$ {}^{L}\hat{\mathbf{x}} = {}^{L}P\,{}^{R}\hat{\mathbf{x}} = {}^{L}P[\mathbf{C}_R\boldsymbol{\Omega} - (\mathbf{C}_R\boldsymbol{\Omega})I]\,{}^{R}P^{+}\,{}^{R}\hat{\mathbf{x}}$$

Equation is of the form ${}^{L}\hat{\mathbf{x}} = H\,{}^{R}\hat{\mathbf{x}}$ and simply states that points on the laser plane induce a homography between coordinates on the left and right image planes, which is consistent with known results [52]. Finally, the error function becomes

$$E = d^2({}^{L}\mathbf{x}, H\,{}^{R}\hat{\mathbf{x}}) + d^2({}^{R}\mathbf{x}, {}^{R}\hat{\mathbf{x}})$$

where $H = {}^{L}P[\mathbf{C}_R\boldsymbol{\Omega} - (\mathbf{C}_R\boldsymbol{\Omega})I]\,{}^{R}P^{+}$. The reconstruction problem can now be formulated as an unconstrained optimization of equation with respect to ${}^{R}\hat{\mathbf{x}}$. Then, the minimum squared error E over all candidates is used to resolve the validation/correspondence problem, and the reconstruction $\mathbf{X}$ .

## 5 Special Case: Rectilinear Stereo and Pin-Hole Cameras

The results of the previous section apply to general camera models and stereo geometry. However, the special case of rectilinear stereo and pin-hole cameras is important as it reduces equation to a single degree of freedom. Furthermore, rectilinear stereo applies without loss of generality (after projective rectification), and the pin- hole model is a good approximation for CCD cameras (after correcting for radial lens distortion). The stereo cameras used in this work are assumed to have unit aspect ra- tio and no skew, and the pin-hole models are parameterized by equal focal length f.

The camera centres $\mathbf{C}_{L,R}$ and projection matrices ${}^{L,R}P$ for rectilinear pin-hole cameras are given by equations . Substituting these into equation , the relationship between the projections of a point on the light plane can be written as:

$$ {}^{C}\Omega = \begin{pmatrix} c_y c_z - s_x s_y s_z - B_0 c_x s_z \\ c_y s_z + s_x s_y c_z + B_0 c_x c_z \\ -c_x s_y + B_0 s_x \\ (s_x s_y s_z - c_y c_z + B_0 c_x s_z) X_0 + (c_x s_y - B_0 s_x) Z_0 + D_0 \end{pmatrix} $$

(3)

In inhomogeneous coordinates, the relationship between ${}^{L}\hat{\mathbf{x}} = ({}^{L}\hat{x}, {}^{L}\hat{y})$ and ${}^{R}\hat{\mathbf{x}} = ({}^{R}\hat{x}, {}^{R}\hat{y})$ given by the homogeneous transformation in equation can be expressed as

$$ {}^{C}\Omega = \begin{pmatrix} \cos\theta_y \\ \theta_x \sin\theta_y + \theta_z \cos\theta_y + B_0 \\ -\sin\theta_y \\ -X_0 \cos\theta_y + Z_0 \sin\theta_y + D_0 \end{pmatrix} $$

(4)

Since the axes of L and R are parallel (rectilinear stereo), the notation $\hat{y} \equiv {}^{L}\hat{y} = {}^{R}\hat{y}$ replaces equation . Rectilinear stereo gives rise to epipolar lines that are parallel to the x-axis, so the validation algorithm need only consider possible correspon- dences on matching scan-lines in the stereo images. Any measurement error in the stripe detection process is assumed to be in the x-direction only, while the y-coordinate is fixed by the height of the scan-line. Thus, the y-coordinate of the optimal projections are also fixed by the scan-line, ie. $\hat{y} = y$, where y is the y-coordinate of the candidate measurements ${}^{L}\mathbf{x}$ and ${}^{R}\mathbf{x}$.
Finally, substituting equations with $\hat{y} = y$ the image plane error E can be expressed as a function of a single variable, ${}^{R}\hat{x}$:

$$ E = ({}^{L}x + \alpha\, {}^{R}\hat{x} + \beta y + \gamma f)^2 + ({}^{R}x - {}^{R}\hat{x})^2 $$

(5)

where the following change of variables is introduced:

$$ \begin{aligned} \alpha &= (Ab - D)/(Ab + D) \\ \beta &= 2Bb/(Ab + D) \\ \gamma &= 2Cb/(Ab + D) \end{aligned} $$

(6)

For the experimental scanner, with $\boldsymbol{\Omega}$ given by equation , $\alpha$, $\beta$ and $\gamma$ can be written as:

$$ \begin{aligned} \alpha &= -\frac{k_1 \cos\theta_y + k_2 \sin\theta_y + k_3}{\cos\theta_y + k_2 \sin\theta_y + k_3} \\ \beta &= \frac{(1 - k_1)(\theta_x \sin\theta_y + \theta_z \cos\theta_y + B_0)}{\cos\theta_y + k_2 \sin\theta_y + k_3} \\ \gamma &= \frac{(k_1 - 1)\sin\theta_y}{\cos\theta_y + k_2 \sin\theta_y + k_3} \end{aligned} $$

(7)

where $\theta_y = me + c$ and the following change of variables is made in the system parameters:

$$ \begin{aligned} k_1 &= -(b + X_0)/(b - X_0) \\ k_2 &= Z_0/(b - X_0) \\ k_3 &= D_0/(b - X_0) \end{aligned} $$

(8)

Optimization of equation now proceeds using standard techniques, setting

$$ \frac{dE}{d\,{}^{R}\hat{x}} = 0 $$

and solving for $\hat{x}$. Let $\hat{x}$ represent the optimal projection resulting in the minimum squared error, $E^*$. It is straightforward to

show (see Appendix B) that the optimal projection is given by

$$R_{\hat{x}^*} = [R_x - \alpha(L_x + \beta y + \gamma f)]/(\alpha^2 + 1) \quad (9)$$

and the minimum squared error $E^*$ for the optimal solution is:

$$E^* = (L_x + \alpha R_x + \beta y + \gamma f)^2/(\alpha^2 + 1) \quad (10)$$

For completion, substituting equation and $R_{\hat{y}^*} = y$ ) gives the corre- sponding optimal projection on the left image plane as

$$L\hat{x}^* = [\alpha 2 L x - \alpha R x - (\beta y + \gamma f)]/(\alpha 2 + 1) \quad (11)$$

Finally, the optimal 3D reconstruction $\mathbf{X}^*$ is recovered by substituting into equation . In non-homogeneous coordinates, the optimal reconstruction leading to the minimum image plane error for candidate measurements $L_{\mathbf{x}}$ and $R_{\mathbf{x}}$

$$\hat{X}^* = \frac{[(\alpha-1)(\alpha^L x - {}^R x) - (\alpha+1)(\beta y+\gamma f)]b}{(\alpha+1)(\alpha^L x - {}^R x) + (\alpha-1)(\beta y+\gamma f)}$$
$$\hat{Y}^* = \frac{2by(\alpha^2+1)}{(\alpha+1)(\alpha^L x - {}^R x) + (\alpha-1)(\beta y+\gamma f)}$$
$$\hat{Z}^* = \frac{2bf(\alpha^2+1)}{(\alpha+1)(\alpha^L x - {}^R x) + (\alpha-1)(\beta y+\gamma f)} \quad (12)$$

The validation problem can now be solved by evaluating $E^*$ in equation for all pairs of candidate measurements on matching scanlines, and selecting the pair with the minimum error (less than some validation threshold). Once the valid measurements have been identified, the position of the light plane is calculated from the encoder count e using equations and the optimal reconstruction is recovered from the image plane measurements .

## 7 Laser Plane Error

The above solution is optimal with respect to the error of image plane measurements, and assumes that the parameters of the laser plane are known exactly. In practice, the encoder measurements are likely to suffer from both random and systematic error due to acquisition delay and quantization. Unlike the image plane error, the encoder error is constant for all stripe measurements in a given frame and thus cannot be minimized independently for candidate measurements on each scan-line.

Let $L_{\mathbf{x_i}}$ and $R_{\mathbf{x_i}}$, $i = 1 \ldots n$ represent valid corresponding measurements of the laser stripe on the n scan-lines in a frame. The reconstruction error $E^*(e)$ for each pair can be treated as a function of the encoder count via the system model in equations . The total error $E^*_{tot}(e)$ over all scan-lines for a given encoder count e is calculated as:

$$E^*_{\text{tot}}(e) = \sum_{i=1}^{n} E^*_i(e) \quad (13)$$

Finally, an optimal estimate of the encoder count $e^*$ is calculated from the minimization

$$e^* = \arg\min_e [E^*_{\text{tot}}(e)] \quad (14)$$

Since $E^*_{tot}(e)$ is a non-linear function, the optimization in equation is implemented numerically using Levenberg-Marquardt (LM) minimization from MIN- PACK , with the measured value of e as the initial estimate.

As noted above, valid corresponding measurements must be identified before calculating $E^*_{tot}(e)$. However, since the correspondences are determined by minimizing $E^*$ over all candidate pairs given the plane parameters, the correspondences are also a function of the encoder count. Thus, the refined estimate $e^*$ may relate to a different set of optimal correspondences than those from which it was calculated.

To resolve this issue, the optimal correspondences and encoder count are calculated recursively. In the first iteration,

correspondences are calculated using the measured encoder value e to yield the initial estimate $e^*$ via equations . The process is repeated until a stable set of correspondences is found.

The above process is applied to each captured frame, and the optimal encoder count $e^*$ and valid correspondences, $^{L}\mathbf{x_i}$ and $^{R}\mathbf{x_i}$ , are substituted into equations to finally recover the optimal 3D profile of the laser.

## 8 Additional Constraints

As already described, robust stripe detection is based on minimization of the image plane error in equation (3.26). However, the minimum image plane error is a nec- essary but insufficient condition for identifying valid stereo measurements. In the practical implementation, two additional constraints are employed to improve the robustness of stripe detection.

The first constraint simply requires stripe candidates to be moving features; a valid measurement must not appear at the same position in previous frames. This is implemented by processing only those pixels with sufficiently large intensity dif- ference between successive frames. While this constraint successfully rejects static stripe-like edges or textures in most scenes, it has little effect on cross-talk or reflections, since these also appear as moving features.

The second constraint is based on the fact that valid measurements only occur within a sub-region of the image plane, depending on the angle of the light plane.

It can be that the inhomogeneous z–coordinate of a single-camera reconstruction $\mathbf{X}$ can be expressed as a function of the image plane projections $^{L,R}\hat{\mathbf{x}}$ a

$$\hat{Z} = \pm \frac{2bf}{(\alpha+1)^{L,R}\hat{x} + \beta y + \gamma f}$$

(15)

where the positive sign is taken for L and negative sign for R. Rearranging the above, the projected x-coordinate of a point on the light plane may be expressed as a function of depth Z and the height y of the scanline:

$$^{L,R}\hat{x} = -\frac{\beta y + \gamma f}{\alpha + 1} \pm \frac{2bf}{\hat{Z}(\alpha+1)}$$

(16)

The extreme boundaries for valid measurements can now be found by taking the limit of equation for points on the light plane near and far from the camera. Taking the limit for distant reflections gives one boundary at:

$$\lim_{\hat{Z}\to\infty} {}^{L,R}\hat{x} = -\frac{\beta y + \gamma f}{\alpha + 1}$$

(17)

Taking the limit $Z \longrightarrow 0$ for close reflections gives the other boundary at $^{L,R}\hat{x} \longrightarrow \pm\infty$. Now, if w is the width of the captured image, valid measurements on a scan-line at height y must be constrained to the x-coordinate ranges

$$^{L}x \in \left[-\frac{\beta v + \gamma f}{\alpha + 1}, +\frac{w}{2}\right]$$

$$^{R}x \in \left[-\frac{w}{2}, -\frac{\beta v + \gamma f}{\alpha + 1}\right]$$

(18)

Stripe extraction is only applied to pixels within the boundaries ; pixels outside these ranges are immediately classified as invalid. In addition to improving robustness, sub-region processing also reduces computational expense by halving the quantity of raw image data and decreasing the number of stripe can-didates tested for correspondence.

## 9 Active Calibration of System Parameters

In this section, determination of the unknown parameters in the model of the light stripe scanner are considered. Let the unknown parameters be represented by the vector

$$\mathbf{p} = (k_1, k_2, k_3, \theta_X, \theta_Z, B_0, m, k)$$

where $k_1$, $k_2$ and $k_3$ were introduced in equations .Since most of the parameters relate to mechanical properties, the straightforward approach to calibra- tion is manual measurement. However, such an approach would be both difficult and increasingly inaccurate as parameters vary through mechanical wear. To overcome this problem, a strategy is now proposed to optimally estimate $\mathbf{p}$ using only image- based measurements of a non-planar but otherwise arbitrary surface with favourable reflectance properties (the requirement of non-planarity is discussed below). This allows calibration to be performed cheaply and during normal operation.

The calibration procedure begins by scanning the stripe across the target and recording the encoder and image plane measurements for each captured frame. Since the system parameters are initially unknown, the validation problem is approximated by recording only the brightest pair of features per scan-line. Let ${}^{L}\mathbf{x}_{ij}$ and ${}^{R}\mathbf{x}_{ij}$, $i = 1 \ldots n_j$, $j = 1 \ldots t$ represent the centroids of the brightest corresponding features on $n_j$ scan-lines of $t$ captured frames, and let $e_j$ represent the measured encoder value for each frame. As described earlier, image plane measurements have independent errors, while the encoder error couples all measurements in a given frame. Thus, optimal system parameters are determined from iterative minimization of the stripe measurement and encoder errors, based on the algorithm first . First, the total image plane error $E^*_{tot}$ is summed over all frames:

$$E^*_{\text{tot}} = \sum_{j=1}^{t} \sum_{i=1}^{n_j} E^*({}^{L}\mathbf{x}_{ij}, {}^{R}\mathbf{x}_{ij}, e_j, \mathbf{p})$$

(19)

where $E^*$ is defined in equation (3.26). The requirement of a non-planar calibration target can now be justified. For a planar target, the stripe appears as a straight line and the image plane measurements obey a linear relationship of the form $x_{ij} = a_j y_{ij} + b_j$. Then, the total error $E^*_{tot}$ reduces to the form

$$E^*_{\text{tot}} = \sum_{j=1}^{t} \sum_{i=1}^{n_j} (A_j y_{ij} + B_j)^2$$

(20)

Clearly, the sign of $A_j$ and $B_j$ cannot be determined from equation (3.39), since the total error remains unchanged after substituting $-A_j$ and $-B_j$. The geometrical in- terpretation of this result is illustrated in Figure 5, which shows the 2D analogue of a planar target scan. For any set of encoder values $e_j$ and collinear points $\mathbf{X}_j$ measured over $t$ captured frames, there exist two symmetrically opposed laser plane generators capable of producing identical results. This ambiguity can be overcome by constraining the calibration target to be non-planar. It may also be possible for cer– tain non-planar targets to produce ambiguous results, but the current implementation assumes that such an object will rarely be encountered.

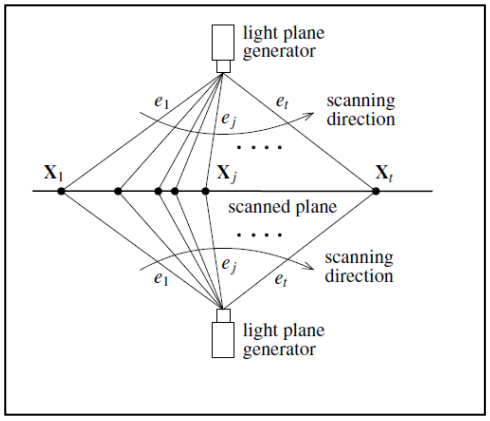


Fig. 5. Possible positions of the light plane from the scan of a planar calibration target. The position of the light plane generator is ambiguous.

An initial estimate $\mathbf{p}^*$ of the parameter vector is given by the minimization

$$\mathbf{p}^*_0 = \arg\min_{\mathbf{p}} [E^*_{\text{tot}}(\mathbf{p})]$$

(21)

using the measured encoder values $e_j$ and stereo correspondences ${}^{L}\mathbf{x}_{ij}$ and ${}^{R}\mathbf{x}_{ij}$. Again, equation is implemented numerically using LM

minimization. The stripe measurements $^{L}\mathbf{x}_{ij}$ and $^{R}\mathbf{x}_{ij}$ are likely to contain gross errors resulting from the initial coarse validation constraint (in the absence of known system parameters). Thus, the next calibration step refines the measurements by applying outlier rejec- tion. Using $e_j$ and the initial estimate $\mathbf{p}^*$, the image plane error $E^*(^{L}\mathbf{x}_{ij}, ^{R}\mathbf{x}_{ij}, e_j, \mathbf{p}^*)$ in equation is calculated for each stereo pair. The measurements are then sorted in order of increasing error, and the top 20% are discarded.

The system parameters and encoder values are then sequentially refined in an iterative process. The initial estimate $\mathbf{p}^*$ is only optimal with respect to image plane error, assuming exact encoder values $e_j$. To account for encoder error, the encoder value is refined for each frame using the method described in Section 3.2.5 with the initial estimate $\mathbf{p}^*$ of the system model. The resulting encoder estimates $e^*$ are optimal with respect to $\mathbf{p}^*$. A refined system model $\mathbf{p}^*$ is then obtained from equation At the $k^{th}$ iteration, the model is considered to have converged when the fractional change in total error $E^*_{tot}$ is less then a threshold $\delta$:

$$\frac{E^*_{\text{tot},k-1} - E^*_{\text{tot},k}}{E^*_{\text{tot},k-1}} < \delta$$

(22)

The final parameter vector $\mathbf{p}^*$ is stored as the near-optimal system model for processing regular.

A final check for global old, based on an estimate of the image plane error. The rare case of non-convergence (less than 10% of trials) is typically due to excessive outliers introduced by the sub- optimal maximum intensity validation constraint applied to the initial measurements. Non-convergence is resolved by repeating the calibration process with a new set of data.

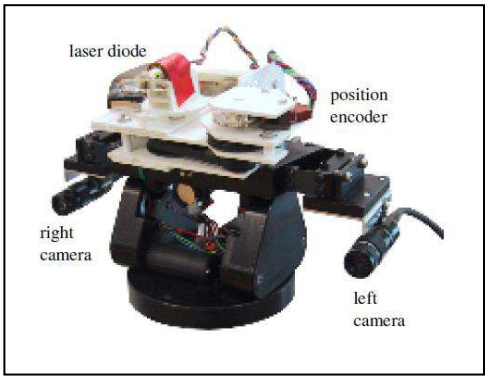


Fig. 6. Experimental stereoscopic light stripe scanner.

The calibration technique presented here is practical, fast and accurate, requiring only a single scan of any suitable non-planar scene. Furthermore, the method does not rely on measurement or estimation of any metric quantities, and so does not require accurate knowledge of camera parameters b and f. Thus, imagebased calibration allows the validation and correspondence problems to be solved robustly and independently of reconstruction accuracy.

## REFERENCES

1 . **B.Adams ,Humanoid robots ,**IEEE Inteligent system **, 2000**

**2. Y Bar Shalom ,Estimation and tracking ,Artech house ,1993**