Substripe Localisation for Improved Structured Light System Performance

Dr Alan M. McIvor and Robert J. Valkenburg

Industrial Research Limited
PO Box 2225, Auckland
New Zealand
E-mail: {a.mcivor,b.valkenburg}@irl.cri.nz

Abstract

Temporally encoded structured light systems are one of the many types of active 3D range sensors available. 3D data is obtained by observing light patterns projected into the scene. In some situations, e.g., calibration, it is necessary to determine the point on the projector emitter plane that corresponds to a given image-plane location, even though the emitter plane is discretised into a finite number of stripes. Thus, a substripe estimator is required. This paper describes three such estimators, and their performance is compared.

Keywords: *3D, substripe accuracy, SLS.*

1. Introduction

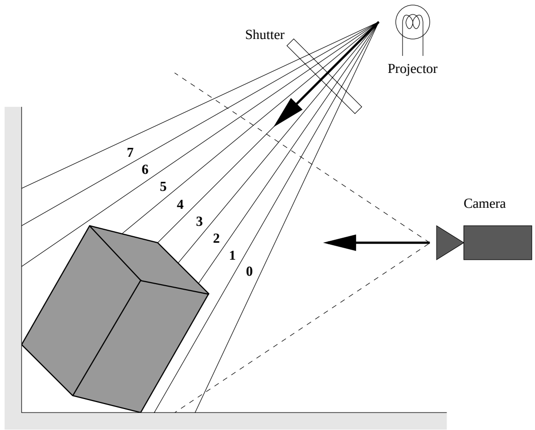
In industrial applications, active sensing methods such as using structured light provide useful techniques for obtaining 3D position information about the visible surfaces in a scene. There are many types of active 3D sensors, as described in [1, 2]. Temporally encoded structured light systems are one such sensor [3, 4, 5]. These use a projector for illuminating the scene with a sequence of patterns, and a camera to acquire data. The observed sequence of patterns is used to determine the correspondence between points on the image plane of the camera and points on the emitter plane of the projector. Subdiscretisation techniques are needed to accurately locate correspondences because of the finite resolution of the camera sensor and the projector emitter. There are several types of subdiscretisation problems [5]. This paper focuses on just one of them: estimating the projector emitter plane coordinates that correspond to given camera image plane coordinates.

The rest of this paper is organised as follows. Section 2 provides some background material explaining how the substripe estimation problem being considered arises. Section 3 describes three methods for doing the substripe estimation and Section 4 evaluates the performance of these methods. Finally, Section 5 provides a conclusion and proposes further work.

2. Background

The form of structured light system being considered in this paper is a temporally encoded structured light system constructed from a camera and a projector, where the projector's emitter plane has been divided into densely packed parallel lines. In an idealised model of such a system, a ray is formed in 3D space by the back-projection of a point on the camera's image plane, i.e., the set of points which project to the given image point. A plane is formed by the projection of a line on the projector's emitter plane. A unique point is identified in 3D space by the spatial intersection of the ray and plane. Projecting light through an emitter line generates an illuminated curve on the visible surfaces, the image of which can be detected in the camera image plane. In this way, the image point and emitter line associated with each visible surface point can be identified. Its 3D coordinates can be obtained by spatial intersection of the associated ray and plane. Using this method, dense 3D surface information can be obtained.

In practice, the lines on the projector's emitter plane will be formed by a shutter with finite thickness stripes (as shown in Figure 1), just as the camera image plane will be formed from finite width and height pixels. Subpixel localisation is used to generate camera image plane coordinate estimates to higher resolution than that of pixels. Similarly, substripe localisation is used to generate projector emitter plane coordinate estimates to higher resolution that that of the stripes. One example of this is to estimate the projector emitter plane coordinate at a given location in the camera image plane. That is, given a point on the camera image plane, its back-projection intersects a visible surface at some point in 3D space. The problem is to estimate which line in the projector's emitter plane would, if light were projected through it, illuminate that point on the surface.



		Line 7	6	5	4	3	2	1	0
Pattern	A	On	On	On	On	Off	Off	Off	Off
	B	Off	Off	On	On	On	On	Off	Off
	C	Off	On	On	Off	Off	On	On	Off

Figure 1: A model of a temporally encoded structured light system showing how the projector emitter plane stripes are encoded over the sequence of images. In the illustration, 8 stripes (labelled $0 \cdots 7$) are encoded by three images (A, B, C).

3. Methods

This paper considers three substripe localisation methods for the problem described in the proceeding section. In what follows, it will be assumed that the emitter plane has been parameterised by a 1D coordinate frame such that the stripe centres have integer coordinates, in analogy to the camera image plane being given a 2D coordinate frame such that the pixel centres have integer coordinates. The integer projector emitter plane coordinate at a stripe's centre will be called the stripe id. The stripe id at any point on the camera image plane can be determined by thresholding the sequence of images observed as the projector pattern is varied, as illustrated in Figure 1.

The three methods are:

3.1 Polynomial fitting

A low-order polynomial is fitted using least squares to the stripe ids observed at the pixel centres in a region of interest centred around the camera image plane point where the substripe estimate is required. Then the substripe estimate is the value of this polynomial at the camera image point in question.

The stripe ids are integer valued and hence they form a function of camera image plane coordinates that is discontinuous, as shown in Figure 2. For small surface patches, only first order polynomials provide reliable estimates. Fitted higher order polynomials are distorted by trying to fit the steps in the stripe id data. For the experiments, first order polynomials were fitted over 17×17 data patches about each point.

3.2 Boundary Interpolation

In an image where a projector emitter plane stripe is illuminated and a neighbouring stripe

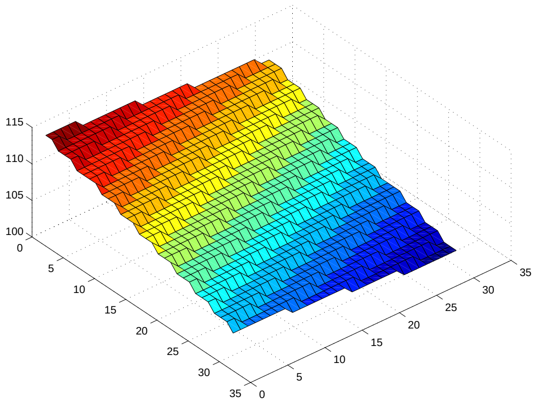
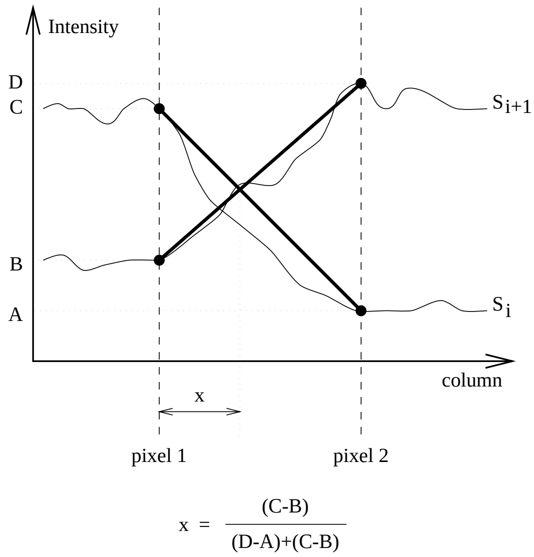


Figure 2: The stripe ids in a 36×36 data patch from a plane, plotted as functions of image plane coordinates. Note the step effect due to the discretisation of the projector emitter plane coordinates.



$$x = \frac{(C-B)}{(D-A)+(C-B)}$$

Figure 3: The operator used for locating zero-crossings in an intensity differential pair. It is applied between two pixels where the difference image changes sign.

is not, an illumination edge is formed by the boundary between the two stripes. This can be localised to subpixel accuracy with an edge detector [6]. The stripe boundaries occur at known emitter plane coordinates. These can be interpolated to estimate the emitter plane coordinate at the given camera image plane point. In the tested implementation, linear interpolation between neighbouring boundaries in a pixel column is used. Since the structured light system used for testing projects the patterns as differential pairs, the illumination edges are detected by looking for zero-crossings in the difference images. A linear zero-crossing localisation operator is used, as shown in Figure 3. This zero-crossing localisation operator has the advantage of using the smallest possible support region, and hence is as insensitive as possible to the effects of the underlying 3D surface. Figure 4 shows the stripe transitions extracted from structured light system data taken from a planar surface, overlaid on top of the pixel grid. The above operator was used to obtain subpixel boundary position estimates. Linear interpolation between stripe boundaries above and below a pixel centre (in the same column) was used to generate projector emitter plane coordinate estimates at each pixel centre.

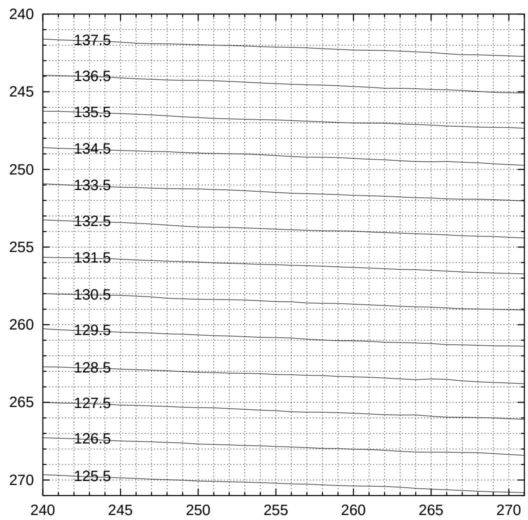


Figure 4: Stripe boundaries detected in the central 32×32 patch of data from a planar surface. The grid connects pixel centres. A boundary's label indicates its the projector emitter plane coordinate.

3.3 Sinusoidal Fitting

In practice, multiple stripes are projected simultaneously, but in such a way that they can still be individually identified over the ensemble of images. In the structured light system used for testing, one such stripe pattern (the band image with the highest spatial frequency pattern) generates intensity variations in the associated image which are sinusoidal in nature, as illustrated in Figure 5. The sinusoidal fitting substripe estimation method involves fitting a spatial sinusoidal model (with 6 harmonics)

$$
\begin{aligned}
s(x,y) &= a_0 + \sum_{k=1}^{6} b_k \cos k(\omega_x x + \omega_y y) \\
&\quad + \sum_{k=1}^{6} c_k \sin k(\omega_x x + \omega_y y) \qquad (1)
\end{aligned}
$$

to the image generated by the highest frequency pattern to determine the phase at the given camera image plane point (see Figure 6). In the above equation, x and y are the camera image plane coordinates and s is the observed projector emitter plane coordinates. The stripe ids are used to perform phase unwrapping to determine the emitter plane coordinate. Fitting the sinusoid is formulated as a separable least squares problem. It is nonlinear in the two parameters ω_x and ω_y. These parameters are solved for using the Levenberg-Marquardt method. The sinusoid was fitted over a 17×17 region of interest around each data point.

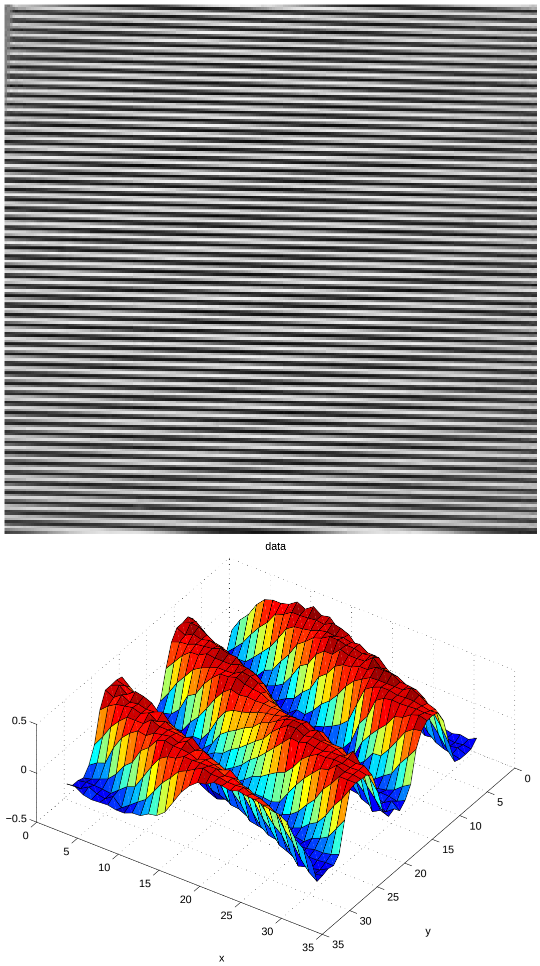


Figure 5: The intensity difference image generated by the highest order band pattern projected onto a planar surface by the Structured Light System. Below this is a plot of the values in a small patch

4. Discussion

The three substripe estimation methods were compared by applying them to data sets obtained from a planar surface. The true projector emitter plane coordinate at each selected image plane location was estimated by fitting a 4th order polynomial model to all of the stripe id data from the plane. The basis for doing this is as follows. The projected patterns are being reflected off a planar surface. If the system is modelled with linear perspective transformation matrices (i.e., pinhole camera and projector models), then the observed projector emitter plane coordinates are a first order rational function of the im-

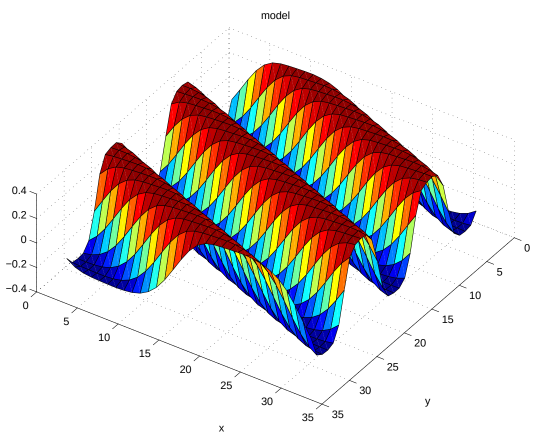


Figure 6: The sinusoidal model fitted to the intensity difference image in Figure 5.

Method	Mean	Std. dev.
Polynomial Fitting	-0.0014	0.0133
Boundary Interpolation	0.0141	0.0089
Sinusoidal Fitting	0.0014	0.0019

Table 1: The statistics of the residual errors in the projector emitter plane coordinate estimates generated by the three substripe localisation methods.

age plane coordinates, where the denominator varies little when the planar surface is almost fronto-parallel, as in this case. This is approximated closely by the 4th order polynomial, which also models the effects of radial distortion, the test surface being non-planar, etc. The residuals from the least square fit are close to being uniformly distributed between -0.5 and 0.5, as expected, indicating that the resulting polynomial is a good model for the observed projector emitter plane coordinates.

The substripe coordinates estimated by the three methods described above were then subtracted from the true projector emitter plane coordinate estimates to generate residual errors. The bias and standard deviation of these are shown in Table 1. These were calculated over the central 32×32 area of the data set. The sinusoidal fitting method produces the best estimates. The planar polynomial fitting method has the largest standard deviation. A factor contributing to this is the changes to the relative position of the support region and stripe boundaries as the support region is shifted from point to point. This causes the fitted plane to tip slightly between data points. The boundary interpolation method has the largest bias. The boundaries are detected as the zero-crossings of the differential pair images. However, it has been observed that these differential pairs, especially the highest spatial frequency ones, can be offset slightly. This would contribute to a bias in the substripe estimates.

The three substripe interpolation methods are very similar in some respects, but differ in their trade-offs between data smoothing and precision of the data extracted for interpolation. Although the polynomial fitting method was described in terms of fitting a polynomial to the stripe ids at the pixel centres, it can also be viewed as fitting a polynomial to the extracted stripe boundaries, where these are only extracted to pixel resolution. The degree of the polynomial makes some implicit assumptions about the surface off which the stripes are being reflected, as the observed projector emitter coordinates are dependent on the shape of this surface. If the surface is planar, then the observed projector emitter coordinates are a first order rational function of the image plane coordinates. Over a small neighbourhood, the terms on the denominator vary little so a linear polynomial is a good approximation. The sinusoidal fitting approach shares these same assumptions about the underlying surface, because it is only under these conditions that the observed high order band image will be periodic. But instead of just using the stripe boundaries extracted to pixel resolution for fitting, all of the high-order band is used. This results in more accurate substripe estimates. In contrast, the boundary interpolation method uses an operator with very small support for localising the stripe boundaries. However, the between boundary interpolation is still closely associated with an implicit planar surface assumption. Whilst the small support means less is assumed about the underlying surface, it also means that the method is more sensitive to noise.

Using substripe interpolation is important when a dense range map (i.e., a range estimate for each camera image pixel) has to be generated. Not only is the data more accurate but it does not suffer from a stair-case effect caused by using the stripe ids alone. Substripe interpolation is also is necessary during calibration. Here, an estimate of the projector emitter plane coordinates at the image plane location of a fiducial mark's image [5]. In this situation, much is also known about the surface on which the fiducial mark's are placed and this can be used in designing a substripe operator. In the system used for testing, the fiducial marks all lie on planar surfaces, and hence the methods discussed above are applicable to this system.

5. Conclusion

One of the problems that occurs when trying to use a temporally encoded structured light system is the substripe interpolation problem. In particular, this problem occurs during system calibration and as a result, accurate substripe estimators are needed if the system is to generate accurate 3D data. One substripe interpolation task is the estimation of the projector emitter plane coordinates that correspond to given image plane coordinates. Three methods were presented for performing this task. The first method is fitting a polynomial to the stripe ids (the discretised version of the projector emitter plane coordinates generated by the stripe labels). The second method was boundary interpolation: extracting the stripe boundaries and then interpolation between then. The third method was fitting a sinusoidal model to the highest order band image projected.

The performance of these three methods was measured by comparing them to best estimates of the true projector emitter plane coordinates for data taken from a planar surface. The polynomial fitting method had the largest standard deviation residual errors. The sinusoidal fitting method had the smallest, almost an order of magnitude better. The boundary interpolation method was inbetween, and had the largest residual bias.

Given that substripe estimation is necessary during system calibration, and thus is needed to generate accurate 3D data, work will continue on improving the performance of these estimators. For example, the boundary interpolation method will be improved by using more complex zero-crossing detection algorithms based on edge detectors to compensate for possible offsets, and spatial interpolation functions with larger support will be used. The methods will also be tested on a variety of non-planar surfaces.

Acknowledgement

This work was funded by the New Zealand Foundation for Research, Science, and Technology under Contract CO8410 "Machine Vision", Objectives 2 "3D Vision" and 4 "Videometrics".

References

[1] R. Jarvis, "Range sensing for computer vision," in *Three-Dimensional Object Recognition Systems* (A. K. Jain and P. J. Flynn, eds.), vol. 1 of *Advances in Image Communications*, pp. 17–56, Amsterdam: Elsevier Science Publishers, 1993.

[2] V. Llario and A. B. Martinez, "Active methods for obtaining depth maps," in *Computer Vision: Theory and Industrial Applications* (C. Torras, ed.), ch. 3, pp. 97–134, Berlin Heidelberg New York: Springer-Verlag, 1992.

[3] K. Sato, H. Yamamoto, and S. Inokuchi, "Tuned range finder for high precision 3D data," in *International Conference on Pattern Recognition*, (Paris, France), pp. 1168–1171, IAPR, IEEE Computer Society, Oct. 1986.

[4] T. G. Stahs and F. M. Wahl, "Fast and robust range data acquisition in a low-cost environment," in *Close-range Photogrammetry meets Machine Vision*, vol. 1395, pp. 496–503, SPIE, 1990.

[5] R. J. Valkenburg and A. M. McIvor, "Accurate 3D measurement using a structured light system," *Image and Vision Computing*, 1997. To appear.

[6] A. M. McIvor, "An alternative interpretation of structured light system data," in *DICTA97*, Massey University, Dec. 1997. To appear (this proceedings).