# An Alternative Interpretation of Structured Light System Data

Dr Alan M. McIvor

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

**Abstract**

*This paper describes an alternative approach to interpreting the data from a temporarily encoded structured light system. Instead of focusing on the stripes generated in the image by such a sensor, the transitions between stripes are of primary importance. Instead of using substripe interpolation to estimate the projector coordinate at image pixel centres and back-projecting these to 3D points, only the transitions are projected. This leads to higher accuracy 3D data.*

***Keywords:*** *3D, structured light.*

## 1. Introduction

One way in which structured light can be used to determine 3D positions in space is to project a plane of light into the scene [1]. The intersection of this plane with the visible surfaces in the scene forms an illuminated stripe on those surfaces. The illuminated stripe can easily be detected in the image plane. The position of the stripe in 3-space (i.e., the position of the illuminated surface points) can be determined by finding the intersection of the projected light plane and the back-projection of the pixels forming the stripe's image. Dense surface data can be obtained by moving the plane through space, e.g., by rotating it about a line (lying in the plane). Such motion can be realised by using a light projector containing a shutter consisting of a number of slits [2, 3]. The light shining through a slit forms a plane. By controlling the slits so that only one is open, and cycling through them all, the effect is as if a light plane is swept through space. The number of images needed to capture the 3D data can be dramatically reduced by having multiple slits open in each image, in such a way that they can be distinguished over the ensemble of images. For example, a slit's opened/closed sequence could be selected on the basis of the 1/0 bits in a grey-code binary encoding of the slit's projector coordinate (where on the projector's image plane the slit is).

This paper describes an alternative way of interpreting the data from a structured light system, which results in improved 3D point estimates. This approach differs from that described above in two fundamental ways. The first, and most important, is that instead of concentrating on the slits per se, the primary features of interest are the transitions between slits. The second difference is that triangulation is performed on just the extracted transitions. This is followed by interpolation in 3D to generate dense position samples if needed.

The rest of this paper is organised as follows. Section 2 provides an overview of the approach and the basis behind it. Section 3 gives details on the transition detection procedure. Section 4 discusses interpolation for providing 3D data at a more dense resolution. Section 5 describes the performance of the system and compares it to previous approaches, and is followed by the conclusion.

## 2. Rationale

The stripe based interpretation of structured light system output can be described as follows: The model of the projector is a 1D coordinate system which identifies the horizontal lines on the projector's image plane. The projection of a line then identifies a plane in space. This intersects a ray coming from the camera's image plane to identify the position of a point in space. This is used to extract the location of points on the curve formed by the intersection of the plane with a surface, since these points are easily identified in the camera image plane (so solving the correspondence problem). In the real system, the lines in the projector's image plane have a finite width. The projection of one of these thus identifies a *wedge* in space. This identifies a finite width ribbon on a surface, commonly called a *stripe*. Thus, to identify the plane coming from the projector which intersects a given ray from the camera (on the surface) requires some substripe interpolation. Substripe interpolation has been identified as one of the most important requirements for good system accuracy [4]. Substripe interpolation is also needed in a laser stripe scanner system. However, in these the stripes have a Gaussian intensity profile, and a simple centroid operator can be used.

When a slit is illuminated but an adjacent slit is not, the projected light pattern has a light to dark transition corresponding to the boundary between the slits. This generates an illumination edge in the resulting images. Such edges are much more like the intersection of a plane with the surface than the stripe generated by a projector slit, as shown in Figure 1. In a real system, the transition between slits will still have finite width but this is significantly less than the width of a slit (In the structured light system used for testing, the transition regions of the shutter span 0.02mm compared to the stripe width of 0.1mm.). The illumination edges can be detected to subpixel accuracy using an edge detector. In the structured light system used for testing, each band pattern (i.e., pattern of on and off slits) is actually projected as a differential pair. Transitions are manifested as zero-crossings in the so-generated difference images. The image plane location of a zero-crossing can be estimated by using linear interpolation between neighbouring pixels in a column which have opposite sign. More complex localisation schemes would use the spatial continuity along a local segment of the transition curve. If the shape of the underlying surface is apriori known, then spatial coherency between neighbouring transitions can also be used as a constraint to improve the accuracy. As well as determining the image plane location of the transitions, it is necessary to identify which transition each detected one is (i.e., determine its projector coordinate). These two aspects are covered in next section.

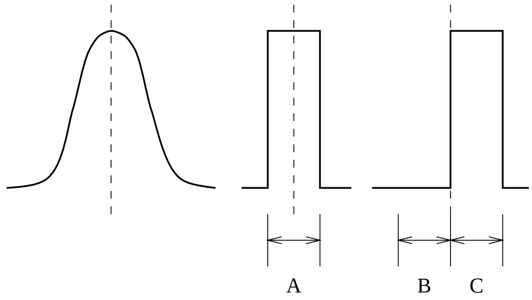


**Figure 1: Three ways of encoding a plane. From left to right: using a laser stripe with a Gaussian profile; using a stripe (A) in a temporally encoded structured light system; using the transition between two stripes (B–C).**

## 3. Transition Detection

A typical transition is shown in Figure 2. The transition between light and dark can be equated directly with the stripe in a laser-striping system. Since the structured light system uses differential pairs of each pattern, the transitions correspond to zero-crossings in the image constructed by taking the difference of the two members of a differential pair. Detecting the zero-crossings is a subpixel locali-

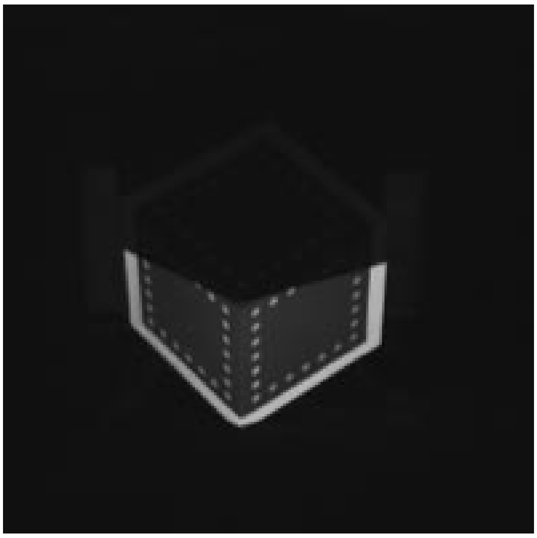

**Figure 2: The first of image pair 2, showing a single transition across the scene.**

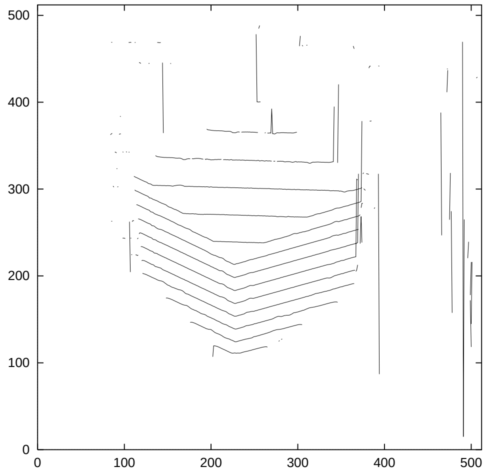


**Figure 3: The transitions between bands in image sequence pair 5 of the cube.**

sation problem. In [5], Sato proposed a simple linear interpolation applied separately to each column that the transition crosses. The transitions so detected in image pair 5 from a cube are shown in Figure 3.

When the band patterns form a grey coding of the stripes' projector coordinates, it is possible to determine the projector stripe coordinate at a transition solely on the basis of the image pair it is in plus the earlier pairs. The only information needed from the earlier pairs is the sign of the difference image at the given camera image plane location. Since no two image pairs have transitions at the same place, the earlier difference images will have a well-defined sign at the location in question, and this is reasonably easy to determine in a robust way.

Image pair $b$ has transitions at stripe locations

$$x_p = 2^{N-b}(1+2i) \qquad i = 0, \ldots, 2^{b-1}-1 \tag{1}$$

where there are $2^N$ lines in the shutter (i.e., $N = 8$ for the K2T shutter since it has 256 lines). The sign of the difference image changes from negative to positive for $i$ even, and from positive to negative for $i$ odd. From this we can determine the sign of an image pair at any stripe coordinate $x_p$. Let

$$j = \left\lfloor \frac{1}{2}\left(\frac{x_p}{2^{N-b}} - 1\right)\right\rfloor \tag{2}$$

then the sign of pair $b$ at $x_p$ is positive if $j$ is even and negative if $j$ is odd.

We can use this to uniquely determine the stripe coordinate of any transition based on the image it is in and the previous images. Image pair 1 has only stripe and this is known to occur at $x_p = 2^{N-1}$. Image pair $b$ has $2^{b-1}$ transitions. When one is detected at a given camera image plane location, each potential stripe coordinate can be considered in turn. For each one, (2) is used to determine the sign each previous pair should have at that stripe coordinate. This is then compared to the actual signs observed at the given camera image plane location. There is only one possible match.

Not all camera pixels will be exposed to a signal, i.e., looking at a surface patch illuminated by the projector, because of effects such as shadowing. The random fluctuations at such pixels can be large enough to result in spurious difference image zero-crossings, which are converted into spurious transitions and hence invalid 3D data. Such outliers need to be detected and rejected to make the data extraction process robust. There are a number of facts that can be used to help detect where unreliable data is being obtained.

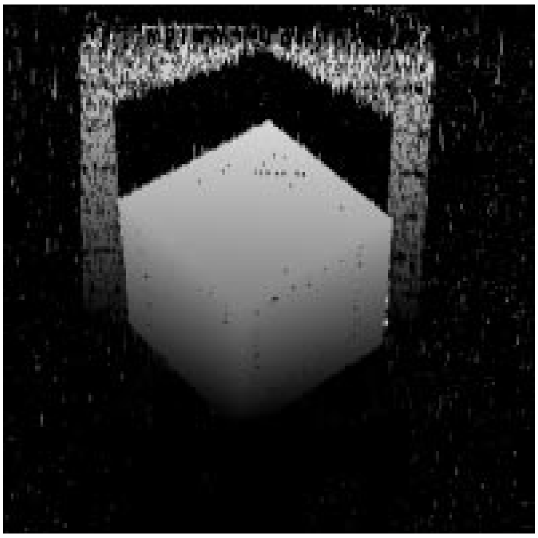

**Figure 4: Dense stripe image constructed by detecting transitions and performing linear interpolation.**



**Figure 5: The stripes projected into 3D. Only every 8th stripe is shown.**

1. On a continuous surface, the observed stripe coordinate (in a camera image plane column) is a continuous function of the camera image plane row coordinate. Hence stripe id changes must be $\pm 1$ [5, Page 659]. However, in practice, the transitions in the highest order images are not reliably detected. Thus, it was found that this constraint was not particularly useful.

2. Once a transition has been identified from an image and the lower order images, the sign of the higher order images at that camera image plane location can be predicted, and these checked for consistency with the observed values. The same comment as above applies.

3. The strength of the zero-crossing can be determined (i.e., the difference between the band images of either side of the transition) and low valued ones rejected.

4. Unless the surface has some particular types of holes, and the position of the sensor is aligned with respect to them, the projector coordinates of the observed transitions should monotonically increase from the bottom of the camera image. Places where this ordering does not occur can be rejected.

## 4. Interpolation

The result of the above image processing is the extraction of the transitions as curves on the image plane (which are continuous except at surface occlusion boundaries). Since each extracted transition's projector coordinates are also known, the extracted image plane curve can be back-projected to a curve in 3-space [6]. Interpolation in 3D between the extracted curves can be performed if denser sampling of the surface is required. This interpolation problem is much simpler than interpolating the projector coordinates in the camera image plane since, for example, a planar surface in 3-space generates a stripe image in which projector coordinates are a rational function of camera image plane coordinates. Furthermore, the interpolation functions represent surfaces and are easier to interpret.

Given the transitions in each image pair, and their projector stripe coordinates, it is possible to generate a dense stripe image by interpolating between the transitions. Figure 4 shows the dense stripe image constructed by linear interpolation between the transitions in each column.

An alternative to generating a dense stripe image is to project the transitions into 3D themselves, is in Figure 5, and do some form of

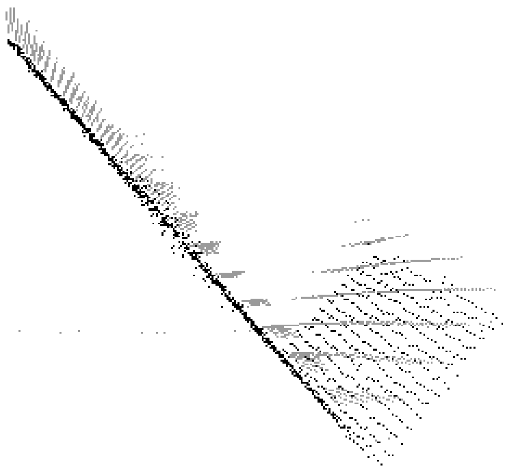

**Figure 6: The 3D data points generated off the planar side of a cube by the original method (grey) and the new method (black).**

| Trial | Stripe | Substripe | Transition |
|---|---|---|---|
| p1 | 2.075780 | 0.739522 | 0.786023 |
| p2 | 1.934560 | 0.695898 | 0.716951 |
| p3 | 1.895870 | 0.643681 | 0.762443 |
| p4 | 3.697820 | 0.646656 | 0.680101 |
| p5 | 1.979860 | 0.469616 | 0.335987 |
| p6 | 4.211060 | 0.486291 | 0.640528 |
| p7 | 4.316820 | 0.549245 | 0.674766 |
| p8 | 1.915020 | 0.658057 | 0.648541 |
| p9 | 1.840910 | 0.435044 | 0.323883 |
| p10 | 1.921660 | 0.773232 | 0.723161 |

**Table 1: Standard Deviation (mm) of offsets from a planar surface fitted to the 3D data generated by the three methods.**

interpolation in 3-space. The interpolation functions can be simpler, since for example, a planar surface in 3-space generates a stripe image in which stripe coordinates are a rational function of camera image plane coordinates. Furthermore, the interpolation functions represent surfaces and are easier to interpret. In some applications, it may not be necessary to generate a dense 3D image. The 3D data along the transitions may be enough.

# 5. Performance

Initial investigation shows that the 3D data so generated is of better quality than that obtained previously, in the sense of being much smoother locally. This is illustrated in Figure 6 which clearly shows the spread of data points off the flat side of a cube is much smaller for this method.

The performance of this method was tested and compared with other methods by applying them to data obtained from a planar surface placed in a number of different poses within the operating volume of the structured light system. Three methods were tested and results are presented in Table 1. The first method (the column labelled "Stripe" in the table) is a stripe detection based method: the stripe incident at a camera pixel is identified but no attempt is made to identify where the pixel is within the stripe. The second method (labelled "Substripe") uses the transition detection approach outlined herein, followed by linear interpolation between transitions (in the columns). This simple form of substripe interpolation is quite accurate when the underlying surface is planar, even though in this case the projector coordinates are actually a first order rational function of camera image plane coordinates. The second method generates a dense set of data like the first, one data point for each camera pixel. In contrast, the third method (labelled "Transition") is the smaller 3D data set generated by spatial intersection from the extracted transitions only. Thus it is a significantly smaller data set. Once a 3D data set had been generated by one of the three methods a planar surface was fitted to the 3D data by total least squares, and the offset of each data point from this surface was calculated. The standard deviation of this offset is given in Table 1. The same system calibration information was used for each method. It clearly shows that the transition based approach generates better data than the stripe detection approach. Since the surface is planar, the simple substripe interpolation method is able to generate dense data as good as using the transitions. However, it won't do so for more complex surfaces, because the interpolation function will not be such a good approximation to the true variation.

## 6. Conclusion

This paper has presented an alternative approach to the interpretation of the raw data generated by a temporally encoded structured light system, and demonstrated that it results in more accurate 3D information. The basis of this approach is concentrating on the transition between stripes and equating them with the planes projected in an ideal structured light system model, instead of using the stripes themselves. The spatial extent of the projector shutter features which generate the transitions is significantly smaller than the stripes themselves.

The transitions are extracted from the images using standard subpixel techniques and, once their identity is established, spatial intersection is used to project them into 3-space. For some applications, 3D positions are needed for the back-projection of camera pixels. These can be obtained by interpolation in 3D between the 3D transition curves, which is a much simpler problem than substripe interpolation in the camera image plane.

This approach results in 3D data which is not only more accurate, but also is smoother locally. Future work will be directed at more accurate transition location detection by using operators similar to the Canny edge detector, and by incorporating the spatial extent of the transition (i.e., the fact that it is a continuous curve on connected surface components). Another aspect of the system that needs further work is the robust rejection of outliers (in shadow regions, etc.).

## Acknowledgement

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

## 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] 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.

[3] 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.

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

[5] K. Sato and S. Inokuchi, "Range-imaging system utilizing nematic liquid crystal mask," in *International Conference on Computer Vision*, pp. 657–661, IEEE Computer Society Press, 1987.

[6] A. M. McIvor and R. J. Valkenburg, "Calibrating a structured light system," in *Image & Vision Computing New Zealand*, (Lincoln, Canterbury), pp. 167–172, Industrial Research Limited, Aug. 1995.