Calibration, Data Consistency and Model Acquisition with a 3-D Laser Striper

E. Trucco
Department of Computing and Electrical Engineering
Heriot-Watt University
Riccarton EH14 4AS, SCOTLAND

R. B. Fisher and A. W. Fitzgibbon
Department of Artificial Intelligence
University of Edinburgh
Edinburgh EH1 2QL, SCOTLAND

D. K. Naidu
MRI Unit, City Hospital
Edinburgh EH10 5SB

Abstract

We analyse the issues of calibration, stripe location and measurement consistency in low-cost, triangulation-based range sensors using structured laser light. We adopt a direct calibration technique which does not require modelling any specific sensor component or phenomena, and therefore is not limited in accuracy by the inability to model error sources. We compare five algorithms for determining the location of the stripe in the images with subpixel accuracy. We describe data consistency tests based on two-camera geometry, which make it possible to acquire satisfactory range images of highly reflective surfaces with holes. Finally, we sketch the use of our range sensor within an automatic system for 3-D model acquisition from multiple range images. Experimental results illustrating the various topics accuracy are reported and discussed.

Paper type: regular paper
Please address all correspondence to: Dr. Emanuele Trucco

Department of Computing and Electrical Engineering
Heriot-Watt University
Riccarton, Edinburgh,
EH1 2QL,
SCOTLAND.
Email: `mtc@cee.hw.ac.uk`
Phone: +44 31 451.3437
FAX: +44 31 451.3431

Footnote for first page

We gratefully acknowledge the invaluable assistance of the staff of our mechanical workshop, directed by Mr. David Wyse. This work was supported by the British Science and Engineering Research Council (now EPSRC), ACME Grant GR/H/86905 “Model Acquisition from Multiple Range Views”.

R. B. Fisher and A. W. Fitzgibbon are with the Department of Artificial Intelligence, University of Edinburgh, Forrest Hill, Edinburgh, EH1 2QL, Scotland, where the work described in this paper was done.
D. K. Naidu is currently with the MRI Unit of the City Hospital, Greenbank Drive, Edinburgh, EH10 5SB, Scotland. E. Trucco is currently with the Department of Computing and Electrical Engineering, Heriot-Watt University, Riccarton, Edinburgh, EH14 4AS, Scotland.

I Introduction

Range images have been attracting increasing attention from the machine vision community, as they directly provide shape descriptions often difficult to compute from intensity images. A variety of physical principles have been applied to the construction of range sensors: extensive reviews can be found in [2, 9, 10]. This paper describes a low-cost range sensor which uses structured laser light to provide a high number of accurate surface measurements per frame. Attention is focused on calibration, data consistency and stripe detection; the application of the sensor in automatic model acquisition from multiple range views is also reported.

The sensor acquires full images (as opposed to profiles) by moving an object on a linear path through a static plane of laser light; the stripe created by the intersection of the object with the laser plane is observed by two opposing cameras, and the pixel coordinates converted into 3-D measurements. The sensor architecture and operation principles are described in Section II. Similar sensors have been reported by several researchers. Saint-Marc *et al.* [27] describe a PC-based scanner achieving an accuracy of 0.25mm (depth) with small objects placed at about 0.5m from the sensor. The laser beam is static as in our case, but the object is moved on a turntable. Comparable arrangements are adopted in [28, 22, 37, 34]. An alternative design leading to more compact sensors can be based on a laser beam directed to sweep a workspace by a system of oscillating and rotating mirrors. Archibald [1] describes the design of such 1- and 2-axis range sensors, which can be small and light enough to be mounted on a robot wrist, and their applications in welding, inspection and NC machining. Compact range cameras have also been used for automatic CAD model acquisition [6], volumetric model reconstruction and visual exploration [36]. Unfortunately the commercial availability of small, light, 2-axis range cameras is still limited. Yet another architecture involving a laser stripe sweeping a scene is proposed by Kanade *et al.* [23]: smart pixels in a fast VLSI sensor record the instant at which the highest intensity caused by a sweeping laser stripe is detected; time measures are converted subsequently into distances.

In the panorama of comparable range sensors, the main contributions of this paper focus on calibration, stripe detection and measurement consistency. We also sketch an advanced application illustrating the use of range images, namely 3-D model acquisition from multiple range views.

The commonly adopted *model-based calibration* estimates the parameters of the geometric transformation that back-projects any points of the image plane of each camera onto the laser plane [2, 9, 27]. This requires a valid closed-form model of the sensor components and basic phenomena, including at least the position, orientation and intrinsic parameters of the cameras, and the position of the light plane. The higher the measurement accuracy required, however, the more phenomena the model must include (e.g. lens distortion, image center, scale factor), thus becoming significantly complicated; some phenomena may always remain elusive. As a consequence, model-based calibration may become unstable and tedious in practice. After some experience with model-based calibration, we devised an alternative method called *direct calibration*, reminiscent of the "black-box" inverse calibration of robotic manipulators [29]. The method starts by measuring the image coordinates of a grid of many known 3-D workspace points, then it builds lookup tables by interpolation, linking pixels to 3-D points. An immediate advantage is that there is no need to model *any* phenomena, since all phenomena are *implicitly* accounted for. The overall accuracy of the method is therefore limited only by the repeatability of the equipment and of the stripe detection algorithm, not by shortcomings of the model. We have designed a simple, fast and automatic procedure implementing direct calibration on our range sensor. The procedure is described in Section III.

In a structured-light triangulation system, the accuracy of 3-D measurements depends on the accuracy of the stripe location in the image. Several methods can be used to achieve high accuracy in stripe location, depending on the accuracy demanded by the application, the type of laser source used, the CCD sensors available and so on. The laser stripe used by Saint-Marc *et al.* [27] is generated using the popular and inexpensive combination of a HeNe laser source and a cylindrical lens, resulting in a Gaussian profile across the stripe. Their observed stripe is wider than 3 pixels; the peak's position is estimated by locating the maximum of a second-order polynomial fitted to the pixel-resolution peak and the two adjacent pixels. Aiming at an accuracy of 1/3,000 over fields of view of 1-30cm depth, with similar depth of field and stand-off distances, Harding [22] uses a diode laser source to generate a "top-hat" stripe profile with sharp edges. In the sensor proposed by Kanade *et al.* [23], individual smart pixels record the instant at which a peak intensity was detected (i.e. the stripe sweeping the scene was imaged) and convert time measurements into distances. Blais and Rioux [5] discuss several peak detection methods for real-time applications, including centre of mass, curve fitting and signal filtering, and propose a FIR filter achieving experimental accuracies of about 0.3 pixel (standard deviation of RMS error). In Section IV we analyse the problem of determining the image position of the peak of a laser stripe with subpixel accuracy. Since various phenomena contribute to distort the ideal Gaussian cross-section of the stripe used in our sensor, it is worth comparing the theoretical and practical characteristics of various subpixel interpolation techniques with those of a fast Gaussian-fitting algorithm.

Another important issue in laser-based range sensing is *measurement consistency*. Many potential applications of range finders are in industrial settings, where objects surfaces are often made of polished metal or plastic. Such surfaces have a specular component, so that noisy reflections of the laser stripe may appear in the images observed by the cameras. These reflections can be mistaken for the primary signal and give rise to false range measurements. For instance, as illustrated in Section V, range images of shiny surfaces with holes may contain spurious peaks or whole surfaces protruding from the holes. In such cases, the data can prove unusable. To obviate this problem, some users of industrial scanners simply coat surfaces with a matte white substance, e.g. tempera paint, which can rinse clean from most parts [39]. This may however be unacceptable, e.g. whenever high accuracies ($100 \mu m$ or less) are required. Section V presents some consistency tests based on two-camera geometry that eliminate most of the spurious range values. The key observation is that specular reflections produce range values depending on camera position. Hence, the range values obtained from each camera can be compared, and points leading to inconsistent range values eliminated.

Our range sensor has been built in the framework of the IMAGINE research project [15, 17, 19], which addresses recognition of complex 3-D objects from range data. In this context, the sensor supports reliable, data-driven surface segmentation and description [32, 33] by providing accurate images even if built with low-cost, off-the-shelf components (a feature of major interest in itself). Indeed the accuracy and repeatibility of our system are currently in the order of 0.15mm (Section VI), a very good result for such a low-cost system (cf. for instance the similar sensor described in [27], with a reported accuracy of 0.25mm).

Experiments and tests are reported and discussed in Section VI. Finally, Section VII sketches the use of the sensor within a recent development of the IMAGINE project aimed at automatic model acquisition from multiple range views [4, 17].

II Sensor architecture

The architecture of our range sensor is sketched in Figure 1. The object to be scanned sits on a platform moved on a linear rail by microstepper motors under computer (Sun3) control through a Compumotor CX interface with in-built RS-232C interface. One microstep is $6.6\mu m$, and the nominal positioning accuracy and repeatability is 2μsteps. Objects must be contained in a parallelepipedal workspace about 15cm each side.

The object is moved through a plane of laser light (output power 2mW at 632.8nm) obtained using a HeNe source, mirrors and a cylindrical lens. The laser light is a beam of circular cross-section with Gaussian intensity profile before passing through the cylindrical lens, which spreads the beam into a thin plane, so that the intensity profile in the workspace of the sensor is practically constant along the stripe and still Gaussian across the stripe. The latter fact is used for determining the position of the stripe in the image with subpixel resolution (Section IV). In practice, the cross-section of the stripe as observed by the cameras is not a perfect Gaussian, as each pixel integrates light over its field of view, the physical sensor pads of the solid-state cameras have gaps between them, the sensor pads have internal structure that affects their sensitivity, and not all sensor pads are equally sensitive.

The planar curve (stripe) resulting from the intersection of the plane of laser light with the object surface is observed by two opposing cameras (off-the-shelf 577x581 Panasonic BL200/B) mounted about one meter from the platform. This camera arrangement limits the occlusion problem and is essential for some of our measurement consistency constraints. The acquired images are stored in a Datacube as 512x512 frames. One millimeter in the vertical direction corresponds to less than one pixel in the images. Several parameters can be controlled by the user, including image scaling factors, depth quantization, image resolution, and the depth interval to be scanned.

III Direct calibration

Our direct calibration method is based on a simple idea. If the camera coordinates of a sufficiently dense grid of workspace points (called *calibration grid*) can be measured accurately, then the position of any point in the workspace can be obtained by inverting the resulting world-to-camera maps and interpolating between surrounding points. thus avoiding the need to model any system component. We have implemented this idea in a two-stage procedure.

A Stage 1: building the calibration grids

In the first step, a calibration grid is built for each camera. We have designed and built a calibration block (Figure 2) consisting of 145 steps, each 2mm in length and 1mm in height. In order to detect calibration points in the Y direction (refer to Figure 1) the block is formed by

20 parallel slices spaced regularly. The only operator intervention required is to place the block on the striper's platform so that the laser stripe falls entirely on the top surface of the lower step. The block is then advanced automatically 2mm at a time, so that the stripe is observed by both cameras on each of the 145 steps. For each position, the stripe appears as a linear sequence of segments (corresponding to the top surfaces of the step's slices), and the position of the segments' centers is detected to subpixel accuracy [25] and recorded. The block's size is designed so that every observable point in the stripe plane lies no farther than 1mm in range (Z axis) and 4mm along the stripe (Y axis) from the nearest calibrated point. The slope of the block allows simultaneous calibration of both cameras. First the y camera coordinate of the centre is located by an adaptive algorithm as the weighted centre of the segment's points. Then the peak of the stripe (x camera coordinate) is obtained using three-pixel Gaussian interpolation (Section IV). An example of the resulting grid of calibration points for the right camera (145 height levels) is shown in Figure 3. The x axis shows the image position of the calibration points, which depends on the height of the block level on which the stripe impinges. The y axis shows the points' image y position. The slight irregularity in the x direction, largely compensated for by second stage 2 (interpolation), is due to irregular changes in the shape of the stripe profile as the stripe sweeps the CCD sensor (addressed in Section C).

B Stage 2: building image-world maps

In the second stage, the calibration grids are inverted and interpolated to obtain a complete look-up table for each camera. Using least-square linear interpolation with 5 or more calibrated points, *each* integer image pixel is associated with a 3-D point within the calibrated workspace.

The purpose of this conversion is threefold. First, the calibration process does not produce calibrated points for every possible observed stripe position. Second, it is too time-consuming to interpolate the neighbouring calibrated points during run-time acquisition. Third, it requires too much storage to record more than the (X, Y, Z) coordinates (as floating point numbers) for more than integer-coordinate pixels. During data acquisition, subpixel values are re-interpolated for intermediate values.

Our direct calibration procedure is simple, automatic and efficient (currently about 15 minutes for 140 height levels with software not optimised for speed).

C Data acquisition

The direct calibration method suggests an efficient algorithm for image acquisition. Once the sensor has been calibrated, a known one-to-one correspondence exists between a 3-D point P on the stripe and its image in each camera, as P is constrained to lie in the plane of laser light. Therefore it is possible to associate unambiguously each stripe point (x, y), in subpixel-precision image coordinates, to a 3-D point (X, Y, Z). This is done simply by using (x, y) to index in the 2-D lookup tables built by the direct calibration. However, as the subpixel image coordinates do not, in general, correspond to the integer table indexes, (X, Y, Z) is obtained by linear interpolation. The dense sampling of the calibrated image space (i.e. the close entries in the lookup tables) supports very good accuracy throughout the sensor's workspace.

Whenever a point is observed by both cameras, two independent measurements (i.e. 3-D points (X_1, Y_1, Z_1) and (X_2, Y_2, Z_2)) are obtained and must be combined into a final measure. First the measurements are checked for consistency as illustrated in Section V. If the measurements are declared compatible, their average is taken as the final measure (which provides, of course, the best *a-priori* estimate in the absence of further statistical information about the measures). If a point on the object surface is visible to one camera only, the final measure is the one obtained from that camera. Surface areas invisible to both cameras are limited by the two-camera arrangement.

Although speed was not a research objective, the acquisition rate can reach a few stripes per second, not a despicable one given the equipment used. Speed could be greatly improved with the use of a synch generator and analogue stripe detection hardware (with a limit of 25x512 points per stripe per second at full video rate). Figure 4 shows a few examples of range images of a well-known casting, acquired by our striper.

IV Locating the stripe with subpixel accuracy

This section compares five algorithms for determining the image position of the peak of the laser stripe to subpixel accuracy, obtained by fitting a 1-D curve to a few pixels around the maximum of the stripe cross-section, which should in theory be Gaussian. In practice, owing to the behaviour of the CCD sensors and to imperfections in the stripe generation equipment, the observed cross-section of the stripe is not Gaussian. Moreover, only a few pixels around the stripe peak are used as support for the fit. For these reasons it is reasonable to ask oneself whether a non-Gaussian fit might result in better performances than fitting a Gaussian one. The algorithms were compared in terms of accuracy, robustness and computational speed. In addition to empirical testing, a theoretical comparison was also devised to provide a framework for analysis of the empirical results, further details of which can be found in [25].

The image is scanned perpendicularly to the width of the observed stripe, i.e. along the x image axis. In the following, when we refer to the subpixel position of the peak of the stripe, we are discussing the x coordinate of the pixel. It is also assumed that the background intensity level (usually about 10) is subtracted from all intensity values before stripe location. The five subpixel algorithms compared are the following.

1. Gaussian approximation. This method fits a Gaussian profile to the three highest, contiguous intensity values around the observed peak of the stripe. If a, b and c are the intensity values observed at pixel positions $i-1$, i and $i+1$ respectively, b being the highest value, then the subpixel location ($\hat{X}$) of the peak is given by :

$$\hat{X} = i - \frac{1}{2}\left(\frac{\ln(c) - \ln(a)}{\ln(a) + \ln(c) - 2\ln(b)}\right)$$

where i is the x pixel coordinate of the centre of the pixel with intensity value b. As a, b and c are integers in the range 0-255, the logarithm can be performed by table lookup. We have not found any previous references to this form of peak detector in the literature.

2. Centre of Mass. The centre-of-mass algorithm computes the location of the peak by weighted average:

$$\hat{X} = \frac{a(i-1) + bi + c(i+1)}{a+b+c} = i + \frac{c-a}{a+b+c}$$

We have considered this algorithm using 3, 5 and 7 points, denoted CoM3, CoM5 and CoM7 henceforth. The extension of the algorithm for the latter two cases is obvious and is omitted. Algorithms using all points along the stripe also exist [38].

3. Linear Interpolation. This method assumes that a simple, linear relationship defines the spread of intensity values across the stripe. Thus, if the three highest intensity values are identified as before:

$$\hat{X} = \begin{cases} x - \frac{(a-c)}{2(b-a)} & c > a \\ x - \frac{(a-c)}{2(b-c)} & otherwise \end{cases}$$

4. Parabolic Estimator

Another peak finder can be derived by Taylor series expansion of the first derivative of the signal intensity near the peak. Let the peak be at $f(i+\delta)$ and assume we observe the signal at $f(i)$; then

$$f'(i+\delta) = 0 = f'(i) + \delta f''(i) + O(\delta^2)$$

Hence, neglecting the higher order terms,

$$\delta \approx -\frac{f'(i)}{f''(i)}$$

And introducing finite difference derivatives:

$$\delta = -\frac{f(i+1) - f(i-1)}{2(f(i+1) - 2f(i) + f(i-1))}$$

As the same peak finder can be found by fitting a parabola to the points $f(i-1)$, $f(i)$ and $f(i+1)$, we call this algorithm *parabolic estimator*.

5. Blais and Rioux Detectors

Blais and Rioux [5] introduced fourth and eighth order linear filters:

$$\begin{aligned} g_4(i) &= f(i-2) + f(i-1) - f(i+1) - f(i+2) \\ g_8(i) &= f(i-4) + f(i-3) + f(i-2) + f(i-1) - \\ & \quad f(i+1) - f(i+2) - f(i+3) - f(i+4) \end{aligned}$$

to which we also add a second order filter:

$$g_2(i) = f(i-1) - f(i+1)$$

These operators act like a form of numerical derivative operator. The estimated peak position is given by

$$\delta = \frac{g(i)}{g(i) - g(i+1)}$$

The results of Blais and Rioux showed that the fourth order operator had better performance than the eight order operator over the stripe widths that we are interested in here, so we only analyze it (called BR4 below) and the simplified second order operator (called BR2 below). The eighth order operator has better performance for stripe widths with Gaussian width parameter larger than 2 pixels. Note that this operator is only applied in the given form for $f(i+1) > f(i-1)$. If $f(i+1) < f(i-1)$, then:

$$\delta = \frac{g(i-1)}{g(i-1) - g(i)} - 1$$

A Maximum Theoretical Error of Estimators

Assume that the true stripe profile is Gaussian and modeled by (up to a multiplicative constant)

$$f(n) = e^{-\frac{(n-\delta)^2}{2\sigma^2}}$$

where $-\frac{1}{2} \leq \delta \leq \frac{1}{2}$ is the true peak position and f is sampled at n = -2, -1, 0, 1, 2. We want to determine the relationship between estimated and true peak positions (i.e. offset) for each of the peak detectors above. We ignore the problem of pixels integrating their inputs over their spatial extent, as well as any shaping functions the camera and digitizer may apply.

In order to estimate the maximum deviation $\mid \delta - \hat{\delta} \mid$ over the range $-\frac{1}{2} \leq \delta \leq \frac{1}{2}$ for each estimator, we calculated $\hat{\delta}$ for synthetic stripe profiles with σ similar to the one of our sensor's real stripe. Then we realised that weighting the estimator ($\hat{\delta}' = \alpha_{estimator}\hat{\delta}$) could, for a given σ, reduce the maximum error by one order of magnitude by spreading the error across the full range. In practice, α must be chosen to maximally reduce the error for a desired σ. The smallest maximum errors achieved for three values of σ and their respective α are given in Table 1. The optimal α for one σ is not necessarily optimal for different σ, and Table 1 shows that only the Gaussian, COM7, BR2 and BR4 algorithms are reasonably insensitive to changes of α.

B Bias of Estimators

In order to determine the bias of the peak estimators, we derived first-order relations linking the estimated peak offset $\hat{\delta}$ to the real offset δ. The relations are summarized in Table 2. Details of the derivation for the Gaussian and linear estimators can be found in Appendix 1. As expected, the Gaussian estimator has ideal form for small δ. When $\sigma = 1.0$, as approximately in our sensor, and making use of α (Section A), the resulting $\hat{\delta}$ is shown below. The conclusion is that, in the ideal, noise-free case, all but the linear and BR2 estimators are reasonably unbiased.

Estimator	Gauss	CoM3	CoM5	CoM7	Linear	Parabolic	BR2	BR4
$\hat{\delta}$	1.00δ	1.01δ	1.00δ	1.00δ	1.40δ	0.83δ	1.33δ	1.17δ

C Errors in the Presence of Noise

In line with the experiments of Blais and Rioux [5], we investigated how additive noise corrupting the stripe affects the estimated stripe position. We generated synthetic stripes with different widths and peaks offset about exact pixel positions by small, random amounts; we then added uniform noise (following the model of Blais and Rioux):

$$s(m, x, \sigma, \beta) = e^{-\frac{(m-x)^2}{2\sigma^2}} + \beta n$$

where $m \in \{-3, -2, -1, 0, 1, 2, 3\}$ is the peak position at pixel resolution, $x \in [-0.5, +0.5]$ is the peak offset, $\sigma \in [0.8, 1.8]$ is the stripe width, $n \in [0, 1]$ is a random variable of uniform density, $\beta \in \{0.0, 0.1, 0.25\}$ is the maximum noise magnitude. Values for β were chosen after observation of noise in our sensor, where the noise is about 1-5% of the peak intensity. We measured the RMS error $\sqrt{\frac{1}{N}\sum(x_i - \hat{x}_i)^2}$ and the maximum deviation $max \mid x_i - \hat{x}_i \mid$ as functions of σ for $N = 10,000$ samples. Figure 5 is a plot of the RMS error as a function of σ for $\beta = 0, 0.1, 0.25$ respectively. Figure 6 show the maximum error for the same values of σ.

The results indicate that CoM3 and CoM5 perform significantly worse than other estimators. The error of the CoM7 estimator at low stripe widths is caused by stripe intensities falling quickly as one moves away from the peak, so that the noise becomes quickly predominant.

We also summed the RMS error for $\sigma = 0.8 - 1.8$ (by 0.05) for the three values of β. The results are summarized in Table 3. The figures suggest good performance over a range of σ and β for BR2 and BR4, which is also clear in Figure 5 ($\beta = 0.1, 0.25$). However, the Gaussian estimator has obvious benefits as the noise level or stripe width decreases.

V Checking data consistency

Figure 7 illustrates the possible effect of spurious reflections of the laser light being mistaken for the true stripe. Two range images of a polished-aluminium block with holes are shown. The images were acquired using two opposing cameras independently (see Section II). Obviously false peaks appear instead of holes. How do reflections from specular surfaces cause such spurious range values?

Figure 8 shows the cross-section of a rectangular or cylindrical hole, taken perpendicularly to the light plane. Suppose that the light stripe is observed after reflection from the specular surface of the hole: the specular reflection at point F is observed rather than the true point T. The false point might be chosen because it is brighter (often possible on specular surfaces) or because the true point is hidden. Since all observed points must lie in the plane of laser light, the height of point Y is incorrectly recorded.

The false range surfaces shown in Figure 7 resulted from this phenomenon occurring at many positions along each of many stripes. The tilting false surface arises because, as the stripe moves away from the wall, the triangulated false position moves further away from the true surface. This simple false-surface pattern arises from the simple rectangular hole geometry; more complex holes or combinations of specular surfaces produce more complex artifacts.

The rejection of false range values is based on the constraints described below. Any points that do not satisfy the constraints are eliminated.

Illumination direction constraint. Assuming that the stripe plane illumination projects from fixed directions (either orthographically or perspectively), it is not possible for a beam of light to intersect the surface twice. Mathematically, each such beam of light projects onto a curve (usually a line) in the the sensor's projection plane. Therefore, the light stripe should intersect this curve in at most one point. When more than one point is observed, all points are eliminated as it is not possible to tell easily which is the correct point (brightness is no guarantee on specular surfaces). Figure 9 illustrates this constraint.

Observable surface constraint. Adjacent stripe positions often lead to nearby noisy points forming spurious range surfaces (Figure 7 shows an example). One constraint that eliminates many such surfaces is the requirement that the visible surface portions must face the observing sensor (otherwise the surface could not have been seen). Figure 10 illustrates this constraint. Hence, any local surface point $\vec{P}_\alpha(t)$ whose normal $\vec{n}_\alpha(t)$ satisfies

$$\vec{n}_\alpha(t) \circ (\vec{P}_\alpha(t) - \vec{O}_\alpha) > 0$$

where $\vec{O}_\alpha$ is the origin of the camera reference frame, should be rejected. This constraint is independent of the number of cameras used.

Consistent surface constraint. If a true point is observed by both cameras, then the range values $Z_L(t)$ and $Z_R(t)$ from both cameras should be the same. If the following condition occurs:

$$| Z_L(t) - Z_R(t) | > \tau_d$$

then the point has been corrupted by spurious reflections and must be rejected. τ_d is chosen based on the noise statistics of true range images.

In addition to having the same Z position, the surface normals of the surfaces observed from the left and right sensors should be the same. Let $\vec{n}_L(t)$ and $\vec{n}_R(t)$ be the local surface normals for the left and right camera data. Then, if the inner product of the normals satisfies

$$\vec{n}_L(t) \circ \vec{n}_R(t) < \tau_n$$

then the point is rejected as corrupted by spurious reflections. τ_n is chosen based on the noise statistics of true range images; however, it may need to be set carefully, since surface normals are related to the first-order derivatives of the data and thus are more affected by noise. Figure 11 illustrates this constraint.

Unobscured-once-viewed constraint. An additional constraint can be derived when two cameras are used. If a point was visible by only one camera, there must be a valid point seen by the other camera that obscures the first point. Hence, any points that are visible to one camera and are not obscured relative to the other camera, yet were not observed, are likely to be spurious and are removed. Figure 12 illustrates this constraint.

VI Experimental assessment

This section reports some of the tests performed in the experimental assessment of our range striper, namely regarding Z accuracy, the effect of the consistency tests, and the experimental evaluation of the peak detector algorithms.

A Accuracy

Table 4 gives the flavour of the accuracy of our striper using direct calibration. The table gives the quantization step δz (using 256 levels), the mean error, e, its standard deviation, σ_e, the mean absolute error, e_a, and its standard deviation, σ_{e_a}, all in mm, measured using both cameras and accurately known planes of different materials and placed at different heights (material and height, in mm, are specified in the leftmost column). Comparable accuracies were obtained using each camera individually. We also noticed that the error magnitude remains constant throughout the field of view, whereas it increased towards the periphery with our previous model-based calibration.

B Data consistency

Figure 13 shows a range image, with consistency tests enforced, of the polished aluminium block with holes which caused the spurious surfaces in Figure 7. The larger holes' diameter is 18mm, the smaller ones' 14mm; depths varied between 9 and 13mm. The dramatic rejection of spurious range values is evident. Some of the true range points have also been eliminated, which has caused a slightly more ragged appearance to the object surface; notice however that the height of the remaining range points has been correctly estimated. In spite of the strong reflections, there are also enough data to estimate the real depth of all holes.

C Evaluating the peak detectors

In order to assess experimentally the accuracy and comparative performance of the various peak detectors, we acquired range images of accurately known planar surfaces, and compared the accuracies of the measurements obtained using the various estimators. Each image consisted of 100 stripes. A small micro-stepper step size (0.2mm) was used to achieve high data density. The depth resolution was kept much smaller than the average expected error (0.1mm) to minimise the effect of quantisation. The best fitting plane (using least-squares) was computed for each surface data set, and the resulting deviations recorded.

In summary, the tests showed that only CoM3 and COM5 perform significantly worse than the other estimators. Any difference between the other algorithms was dominated, in our equipment, by the magnitude of the periodic oscillation shown in Figure 14. This was due to variations in the CCD sensor geometry and to the different sampling frequencies of the CCD sensor and the frame buffer used. Consequently, the spatial frequency of the oscillation depended on the slope of the plane observed with respect to the camera viewing direction (Figure 14). Notice that this oscillation make the calibration grids slightly irregular, and this phenomenon can be observed in the example shown in Figure 3; however, the interpolation involved by the direct calibration seems to compensate satisfactorily for this irregularity.

Having no control on the above parameters with our equipment, we limited the effect of the oscillation by using the Gaussian peak detector with samples spaced by 2 pixels (i.e. considering g_{i-2}, g_i, g_{i+2} instead of g_{i-1}, g_i, g_{i+1}, where i is the position of the stripe maximum at pixel precision). This reduced consistently the spatial frequency of the oscillation (more than 50%), and was the algorithm finally adopted in our sensor. The Gaussian estimator introduces also benefits as the noise level or stripe width decreases. Making the sensor achieve higher accuracies would require a closer look at the CCD geometry as well as sampling frequencies.

VII Range-based 3-D model acquisition

We conclude this paper with an example of how full-frame range images can be used in a *automatic model acquisition system*, i.e. a system capable of generating automatically a CAD-like model of an object from a number of images acquired from different viewpoints in space. There is an increasing interest of researchers for this technology [7, 8, 12, 24, 26, 30, 31] as the benefits reliable model acquisition are expected in diverse areas, e.g. reverse engineering, product styling, flexible NC machining, classification, recognition and inspection [39].

The two essential issues that any model acquisition system must address are *model represen-*

tation and *view registration.* Several representations can be in principle adopted to express the models: splines, surface patches, triangulations, volumetric descriptions and finite element models are examples of possible choices. The choice of a representation is in turn linked intimately to the problem of estimating the transformation registering two successive views of the object.

In our system we use two complementary model representations, each of which implies a different solution to the registration problem: a conventional, symbolic surface patch representation [11, 14, 32] is combined with a B-spline model. An example of the two representations, obtained for *a single view* of the casting shown in Figure 4, is given in Figure 15 (left: patches, right: splines). The symbolic model allows fast indexing from a large database, quick pose estimation (due to the small number of corresponding features), and easy specification of inspection scripts (for example, the system can be instructed to "measure diameter of hole 2 in patch B"). On the other hand, pose estimation is limited in accuracy by the small number of correspondences and errors in the surface fitting, and provides only an incomplete surface coverage: only the most stable surface patches are retained in the model. This lack of complete data is undesirable for reverse engineering tasks, and is cured by the use of spline models. Using these models, initial pose estimates can be optimised (albeit expensively), and complete surface models obtained. The following sections sketch the functionalities of our system.

A Model construction from a single range view

First depth and orientation discontinuities are detected from the raw range data, and used as *a priori* weights for diffusion smoothing [32]. Mean and Gaussian curvatures are then calculated from the smoothed image, and the data divided into homogeneous regions according to the sign of the curvatures. Each region is then approximated by a viewpoint-invariant biquadratic patch [21] and finally expanded to include neighboring pixels which are within 3σ ($\sigma = 0.15mm$) of the fitted surface. After this segmentation stage, the region boundaries are approximated by polylines and conics, and adjacent regions are intersected to produce more consistent boundaries. The resulting description is converted into a vision-oriented modeling representation, the Suggestive Modelling System or SMS [13, 20] for visualization and use in our model matching module.

The spline model is constructed by laying a regular grid of control points on the image and fitting a third-order B-spline to the data. Background and noise points are removed in advance. The density of the grid is currently determined by the required accuracy – a 50x50 sampling allows the object in Figure 15 (right) to be approximated to within a maximum error of 0.8mm. An obvious extension is to allocate the knot points more densely at regions of high curvature, as the curvature maps are available from the segmentation process. We plan to implement this in the near future, and expect a significance reduction in error.

B Estimating the transform between views

In order to estimate the parameters of the rigid transformation which relates two views of an object, we assume that the images overlap and start by applying the segmentation process described above to each image, thus producing two lists of surface patches. From these, an interpretation tree algorithm [16] finds consistent sets of corresponding pairs of surfaces. The pairs allows us to compute the 3-D pose of the object using least-squares techniques. The pose is used as an initial estimate for an iterated extended Kalman filter, which computes the uncertainty of the pose estimate [35].

The accuracy of view registration is within about 1σ of the noise on the range data if three or more linearly independent planar surfaces can be extracted reliably from the object (Figure 16). Failing that, biquadratic surfaces estimated about the patch centroids are used to constrain the pose and then translation accuracy falls to about 5mm. If the pose needs to be constrained by using paired centroids, the system is open to error due to occlusion. The rotational accuracy of registration is generally within 1 degree.

C Refining the inter-view transform

Given an initial pose estimate from the symbolic model matcher, the spline model can be used to refine the estimate. The pose is optimized using the Iterated Closest Point (ICP) algorithm [3]. Tests with 2-D examples suggest that the region of convergence occupies about 25% of the space of all possible poses. A less complete investigation with the object above indicates convergence with up to 90 degrees of initial rotation error. The disadvantage of this technique is its computational complexity: for each data point, we must locate the nearest point on the model surface, then calculate the registering transform. Locating the closest point is sped up by a combination of multigridding and subsampling the basic gradient descent algorithm. The registration accuracy is "optimal" in the sense that the noise statistics of the residuals are symmetric and white. Non-convergence does occur however, and we are currently investigating ways of further correcting for this.

D View registration and model postprocessing

Final processing on the models includes merging the single-view descriptions into a single reference frame. This is done easily thanks to the SMS representation for surface patches, which separates patch descriptions into shape, extent and position. The spline models may be treated similarly, by calculating a new fitting spline for the merged sets of range data. As an example of the results, Figure 17 shows three views of a (partial) SMS model, in object-centred coordinates, acquired automatically from several unregistered, overlapping range images of the metal casting shown in Figure 4, which shows that features not visible simultaneously are included and aligned correctly in the model.

VIII Conclusions

We have addressed several issues concerning the design of triangulation-based range sensors using structured illumination, namely a complete direct calibration procedure; consistency constraints to improve the quality of measurements with reflective objects; and a comparative analysis of five algorithms for detecting the peak of the laser stripe at subpixel accuracy. As an example of an advanced machine vision application based on our sensor, we have sketched a 3-D model acquisition system developed in our laboratory. Although our sensor was built with off-the-shelf, low-cost components, the performance achieved is very satisfactory and the application supported adequately.

We believe this paper offers the following contributions. First, calibrating complex sensors can be complex for the designer and tedious for the user. Closed-form models on which calibration is usually based may grow very complicated in order to include all phenomena. Our method for direct calibration of small-workspace sensors is not limited by model inadequacies, and proves simple, practical, and capable of producing very good accuracy.

Second, the consistency tests described can improve dramatically range measurements in the presence of highly reflective surfaces and holes, and eliminate most of the wrong measurements arising from spurious reflections. This problem is usually circumvented in applications by coating reflective objects with matte paint, but this is obviously not always possible. Our solution can be of considerable interest for 3-D shape inspection applications.

Third, given the increasing popularity of range image sensing, the results of our comparative analysis of subpixel stripe detectors can hopefully be of use to others.

Fourth, we have illustrated the use of our range sensor for automatic model acquisition of shape models from multiple range images, a topic of high applicative interest. Our feature-based SMS models are meant primarily for use in 3-D shape-based vision systems, but accurate reconstruction of surfaces for inspection purposes is also being pursued.

References

[1] C. Archibald, "Robot Wrist-Mounted Laser Range Finders and Applications", Proc. Electronic Imaging '90 East Exposition and Conference, Boston, Mass., 1990.

[2] P. Besl, "Active, Optical Imaging Sensors", Machine Vision and Applications, 1988, pp. 127 – 152.

[3] P. J. Besl and N. Mackay, "A Method for Registration of 3-D Shapes", IEEE Trans. Patt. Anal. Mach. Intell., vol. PAMI-14, 1992, pp. 239 – 256.

[4] E. Bispo, A. W. Fitzgibbon and R. Fisher, "Visually Salient 3-D Model Acquisition from Range Data", Proc. British Mach. Vis. Conf. BMVC-93, Surrey, 1993, pp. 619–628.

[5] F. Blais and M. Rioux, "Real-Time Numerical Peak Detector", Signal Processing 11, 1986, pp 145-155.

[6] C. Bradley and G. W. Vivkers, "Free-form Surface Reconstruction for Machine Vision Rapid Prototyping", Optical Engineering 32 (9), 1993, pp. 2191 – 2200.

[7] Y. Chen and G. Medioni, "Object Modeling by Registration of Multiple Range Images", Image and Vision Computing, vol. 10, no. 3, 1992, pp. 145 – 155. pp. 2724 – 2729.

[8] V. Cappellini, R. Casini, M.T. Pareschi and C. Raspollini, "From Multiple Views to Object Recognition", IEEE Transactions on Circuits and Systems, vol. CAS-34, 1987, pp. 1344 – 1350.

[9] H. R. Everett, "Survey of Collision Avoidance and Ranging Sensors for Mobile Robots", Robotics and Autonomous Systems (5), 1989, pp. 5 – 67.

[10] H. R. Everett, D. E. DeMuth and E. H. Stitz, "Survey of Collision Avoidance and Ranging Sensors for Mobile Robots", Revision 1, US Navy Naval Ocean Systems Center, San Diego (CA), 1992.

[11] T.-J. Fan, "Describing and Recognizing 3-D Objects Using Surface Properties", Springer-Verlag, Berlin, 1990.

[12] F. Ferrie, J. Lagarde and P. Waite, "Recovery of Volumetric Object Descriptions from Laser Rangefinder Images", Proc. 1st European Conf. on Comp. Vision (ECCV-90), 1990.

[13] R. Fisher, "SMS: A Suggestive Modelling System for Object Recognition", Image and Vision Computing, vol. 5, 1987, pp. 98 – 104.

[14] R. Fisher, "From Surfaces to Objects", John Wiley, London, 1989.

[15] R. Fisher, "The Design of the IMAGINE2 Scene Analysis Program", in *3D Model Recognition from Stereoscopic Cues*, J. Mayhew and J. Frisby (eds.), MIT Press, 1991.

[16] R. Fisher, "Performance Evaluation of Ten Variations of the Interpretation Tree Matching Algorithm", Proc. Europ. Conf. on Comp. Vis. ECCV-94, Stockholm, 1994.

[17] R. Fisher, A. W. Fitzgibbon, M. Waite, E. Trucco, M. J. L. Orr, "Recognition of Complex 3-D Objects from Range Data", Proc. 7th IAPR Int. Conf. on Image Anal. and Proc., Bari, 1993, pp. 599 – 606.

[18] R. Fisher, D.K. Naidu and D. Singhal, "Rejection of Spurious Reflections in Structured Illumination Range Finders", *Proc. 2nd Conf. on Optical 3-D Measurement Techniques*, Zurich, 1993.

[19] R. Fisher, E. Trucco, M. Waite, A. Fitzgibbon, M. Orr, "IMAGINE: a 3-D Vision System", Sensor Review (2), no. 13, April 1993, pp. 23 – 26.

[20] A. W. Fitzgibbon, "Suggestive Modelling for Machine Vision", Proc. SPIE-1830 Int. Conf. on Curves and Surfaces in Comp. Vis. and Graphics, Boston, 1992, pp. 315 – 320.

[21] A. W. Fitzgibbon and R. Fisher, "Invariant Fitting of Single-Extremum Surfaces", Proc. British Machine Vision Conf., 1993, pp. 569 – 578.

[22] K. Harding, "Tolerance Design of a High Precision Optical Gage", Proc. SPIE Conf. on Optomech. and Electro-Opt. Des. and Ind. Syst., vol. 959, 1988, pp. 247 – 265.

[23] T. Kanade, A. Gruss and L. R. Carley, "A VLSI Sensor Based Rangefinder System", Proc. 5th Int. Symp. on Rob. Res., 1989.

[24] V. Koivunen and J.M. Vezien and R. Bajcsy, "Procedural CAD Models from Range Data", Towards World Class Manufacturing'93, Phoenix, Arizona (USA), 1993.

[25] D. K. Naidu and R. Fisher, "A Comparative Analysis of Algorithms for Determining the Peak Position of a Stripe to Subpixel Accuracy", Proc. British Machine Vision Conf., 1991, pp. 217 – 225.

[26] B. Parvin and G. Medioni, "B-reps from Unregistered Multiple Range Images", Proc. IEEE Int. Conf. on Robotics and Automation, 1992, pp. 1602 – 1607.

[27] P. Saint-Marc, J.-L. Jezouin and G. Medioni, "A Versatile PC-Based Range Finding System", IEEE Trans. Robot. Autom., vol. RA-7, no. 2, April 1991, pp. 250 –256.

[28] Y. Sato, H. Kitagawa, H. Fujita, "Shape measurement of curved objects using multiple slit-ray projections", IEEE Trans. Patt. Anal. Mach. Intell., vol. PAMI-4, no. 6, 1982, pp. 641 – 646.

[29] J. Shamma and D. Whitney, "A Method for Inverse Robot Calibration", ASME Journ. of Dyn. Syst., Meas., and Control (109), March 1987, pp. 36 – 43.

[30] S. Sinha, K. Skifstad and R. Jain, "From 3-D Imaging to CAD: Bridging the Gap to Growth", Advanced Imaging, Nov. 1992, pp. 42 – 45.

[31] M. Soucy and D. Laurendau, "Multi-Resolution Surface Modeling from Multiple Range Views", Proc. IEEE Conf. on Comp. Vision and Patt. Recogn. (CVPR-92), 1992.

[32] E. Trucco and R. Fisher, "Computing Surface-Based Representations from Range Images", Proc. IEEE Int. Symp. on Intell. Control, Glasgow, 1992, pp. 275 – 280.

[33] E. Trucco and R. Fisher, "Experiments in Curvature-Based Segmentation of Range Data", to appear in IEEE Trans. Patt. Anal. Mach. Intell., 1995.

[34] B. Vemuri and J. K. Aggarwal, "Representation and Recognition of Objects from Dense Range Maps", IEEE Trans. Circuits, Man and Systems, vol. CAS-34, no. 11, 1987, pp. 1351 – 1363.

[35] M. Waite, M.J.L. Orr, R.B. Fisher, "Statistical Partial Constraints for 3-D Model Matching and Pose Estimation Problems", Proc. BMVC93 British Machine Vision Association Conference, Surrey, September, 1993.

[36] P. Whaite and F. Ferrie, "From Uncertainty to Visual Exploration", IEEE Trans. Patt. Anal. Mach. Intell., vol. PAMI-13, no. 10, 1991, pp. 1038 – 1049.

[37] S. J. White, "High Speed Scanning Method and Apparatus", U.S. Patent No. 4,498,778, Technical Arts Corp., 1985.

[38] S. J. White, "Method and Apparatus for Locating Center of Reference Pulse in a Measurement System", U.S.Patent No. 4,628,469, December 1986.

[39] T. Wohlers, "3-D Digitizers", Computer Graphics World, July 1992, pp. 73 – 77.

Appendix 1: derivation of bias expressions for peak estimators

Consider a first-order approximation of the Gaussian stripe model of Section IV, and assume that $|\delta|$ is small:

$$f(n) = e^{-\frac{(n-\delta)^2}{2\sigma^2}} \cong (1 + \frac{n\delta}{\sigma^2})e^{-\frac{n^2}{2\sigma^2}}$$

Then, for the **Gaussian** estimator:

$$\begin{aligned}
\hat{\delta} &= \alpha\frac{1}{2}\frac{log(f(-1)) - log(f(1))}{log(f(-1)) + log(f(1)) - 2log(f(0))} \\
&\cong \alpha\frac{1}{2}\frac{log(e^{-\frac{1}{2\sigma^2}}(1-\frac{\delta}{\sigma^2})) - log(e^{-\frac{1}{2\sigma^2}}(1+\frac{\delta}{\sigma^2}))}{log(e^{-\frac{1}{2\sigma^2}}(1-\frac{\delta}{\sigma^2})) + log(e^{-\frac{1}{2\sigma^2}}(1+\frac{\delta}{\sigma^2})) - 2log(1)} \\
&= \alpha\frac{1}{2}\frac{log(1-\frac{\delta}{\sigma^2}) - log(1+\frac{\delta}{\sigma^2})}{2log(e^{-\frac{1}{2\sigma^2}}) + log(1-\frac{\delta}{\sigma^2}) + log(1+\frac{\delta}{\sigma^2})}
\end{aligned}$$

where α is the weight introduced in Section IV. Then, using $log(1 \pm a) \cong \pm a$, we have

$$\begin{aligned}
\hat{\delta} &\cong \alpha\frac{1}{2}\frac{-\frac{\delta}{\sigma^2} - \frac{\delta}{\sigma^2}}{2(-\frac{1}{2\sigma^2}) - \frac{\delta}{\sigma^2} + \frac{\delta}{\sigma^2}} \\
&= \alpha\frac{1}{2}\frac{-2\frac{\delta}{\sigma^2}}{-\frac{1}{\sigma^2}} \\
&= \alpha\delta
\end{aligned}$$

For the **Linear** estimator:

$$\begin{aligned}
\hat{\delta} &= \alpha\frac{f(1) - f(-1)}{2(f(0) - f(-1))} \\
&\cong \alpha\frac{e^{-\frac{1}{2\sigma^2}}(1+\frac{\delta}{\sigma^2}) - e^{-\frac{1}{2\sigma^2}}(1-\frac{\delta}{\sigma^2})}{2(1 - e^{-\frac{1}{2\sigma^2}}(1-\frac{\delta}{\sigma^2}))} \\
&\cong \alpha\frac{e^{-\frac{1}{2\sigma^2}}(2\frac{\delta}{\sigma^2})}{2(1 - e^{-\frac{1}{2\sigma^2}})} \\
&= \alpha\frac{\delta}{\sigma^2}\frac{e^{-\frac{1}{2\sigma^2}}}{(1 - e^{-\frac{1}{2\sigma^2}})}
\end{aligned}$$

σ	Gaussian	CoM3	CoM5	CoM7	Linear	Parabolic	BR2	BR4
0.5	0.0	0.380	0.041	0.021	0.103	0.156	0.026	0.023
1.0	0.0	0.005	0.002	0.000	0.030	0.029	0.024	0.013
1.5	0.0	0.239	0.150	0.057	0.049	0.034	0.022	0.011
α	1.0	1.85	1.093	1.006	0.93	1.08	0.95	0.975

β	Gaussian	CoM3	CoM5	CoM7	Linear	Parabolic	BR2	BR4
0.00	0.00	3.71	1.36	0.31	0.87	0.49	0.39	0.24
0.10	1.07	3.90	1.86	1.32	1.36	1.23	0.93	0.77
0.25	2.49	4.25	2.67	2.63	2.62	2.61	2.12	1.86

surface type	δz	e	σ_e	e_a	σ_{e_a}
matt black,50	0.0013	-0.160	0.095	0.167	0.095
red paint,50	0.0043	0.126	0.106	0.154	0.061
anodized black,100	0.0236	-0.096	0.079	0.100	0.074
polished alum.,100	0.0029	0.066	0.147	0.132	0.093

Estimator	Local Estimate	Estimator	Local Estimate
Gaussian	$\alpha\delta$	CoM3	$\alpha\frac{2\delta}{\sigma^2}\frac{e^{-\frac{1}{2\sigma^2}}}{(1+2e^{-\frac{4}{2\sigma^2}})}$
Linear	$\alpha\frac{\delta}{\sigma^2}\frac{e^{-\frac{1}{2\sigma^2}}}{(1-e^{-\frac{1}{2\sigma^2}})}$	CoM5	$\alpha\frac{2\delta}{\sigma^2}\frac{e^{-\frac{1}{2\sigma^2}}+4e^{-\frac{4}{2\sigma^2}}}{(1+2e^{-\frac{4}{2\sigma^2}}+2e^{-\frac{4}{2\sigma^2}})}$
Parabolic	$\alpha\frac{\delta}{2\sigma^2}\frac{e^{-\frac{1}{2\sigma^2}}}{(1-e^{-\frac{1}{2\sigma^2}})}$	CoM7	$\alpha\frac{2\delta}{\sigma^2}\frac{e^{-\frac{1}{2\sigma^2}}+4e^{-\frac{4}{2\sigma^2}}+9e^{-\frac{9}{2\sigma^2}}}{(1+2e^{-\frac{1}{2\sigma^2}}+2e^{-\frac{4}{2\sigma^2}}+2e^{-\frac{9}{2\sigma^2}})}$
BR2	$\alpha\frac{2\delta}{\sigma^2}\frac{e^{-\frac{1}{2\sigma^2}}}{(1-e^{-\frac{4}{2\sigma^2}})}$	BR4	$\alpha\frac{2\delta}{\sigma^2}\frac{e^{-\frac{1}{2\sigma^2}}+2e^{-\frac{4}{2\sigma^2}}}{(1+e^{-\frac{1}{2\sigma^2}}-e^{-\frac{4}{2\sigma^2}}-e^{-\frac{9}{2\sigma^2}})}$