AUTOMATED SURFACE EXTRACTION IN REAL TIME PHOTOGRAMMETRY

Ming ZHOU*, Clive S. FRASER**
***University of Glasgow, United Kingdom
3D-MATIC Faraday Partnership, Department of Computing Science
ming@dcs.gla.ac.uk
**University of Melbourne, Australia
Department of Geomatics
c.fraser@engineering.unimelb.edu.au

Working Group V/1

KEY WORDS: Algorithms, Multi-sensor, Real-time, Surface Reconstruction

ABSTRACT

Considerable research attention has been given to close-range videometric systems for vision-based, three-dimensional object measurement over recent years. These systems have to demonstrate high speed, high accuracy capability and high flexibility to compete as a three-dimensional coordinate measurement tool. Especially in industries where instant feedback is required, such as production line or tracking applications which demand a process control response, real-time or near-real-time photogrammetric systems are needed. Automatic feature correspondence determination is the central problem in the development of such systems.
This paper reports on the development of a multi-sensor digital close-range photogrammetric (videometric) system for the determination of surface contours of featureless and targetless objects. The use of structured light projection in such systems is a common approach for providing necessary surface texture to support image matching. A new strategy which involves the hierarchical use of projected non-repeating patterns to enable fast, unambiguous image correspondences to be established in an initial coarse surface extraction has been employed in this system. The final, refined surface contours are then recovered via feature-based matching to a projected 'random' pattern.
The paper first describes the non-repeating pattern strategy and the automated image matching approach for surface reconstruction. This is followed by a description of the multi-sensor vision metrology system built to facilitate the experimental work of this research. The integration of the approaches into the system is discussed and experimental measurements conducted are summarised.

1 INTRODUCTION

Digital close-range photogrammetric systems can be designed and developed to comply with the required conditions and purpose of different types of applications. One of the principal factors limiting the application of automated multi-sensor, real-time digital photogrammetry for surface contour determination in industrial measurement is the need for textured object surfaces to support image matching. The use of structural light has been a common approach for provision of the necessary texture on otherwise featureless surfaces. Typical projection patterns employed to date include regular grids, point clouds, lines and random patterns. For example, Jeschke (1990) has applied a regular dense grid of structured light to help photogrammetrically measure the surface profile of soil, rock and cement; and Trinder et al. (1990) have employed a projected grid for the measurement of a mannequin. In their automated system for the all-round mapping of a human body, Wong et al. (1992) have used a projected dot pattern to provide texture, whereas Grün and Baltsavias (1989) have presented a method employing projected vertical lines for the measurement of a human face. Of the patterns mentioned, the random pattern approach usually provides the best opportunity for obtaining high surface point density. Normally, these patterns require interactive input during the procedure of pattern recognition. Some of the pattern projections, which support automatic correlation, give rise to high computational expense. This slows what should desirably be a real-time surface extraction process in a multi-sensor vision metrology system. The use of automatically recognisable projected patterns is therefore preferable.

In order to overcome difficulties in image matching, a hierarchical approach is often adopted whereby the matching process is successively refined using image pyramids (Ackerman & Hahn, 1991). Matching commences with the lowest resolution, resampled imagery, with the results being used at successively higher resolutions to support the determination of search windows for more precise feature correspondence determination. Thus, by the time image matching occurs at the base, full resolution level of the pyramid, a sufficiently refined approximate digital surface model (DSM) will already be at hand.

This paper discusses an alternative hierarchical surface contour determination approach for on-line, multi-sensor vision metrology systems (VMS), namely the use of projected non-repeating patterns in which every point can be unambiguously identified. Thus, object point correspondences can be determined without recourse to image matching *per se*. The purpose of employing non-repeating patterns of successively higher point density is to provide a digital surface model of sufficient definition to support a final image matching stage in which the ultimate surface contour is recovered. The non-repeating pattern (NRP) approach is not considered as a means to replace either feature-based or intensity-based matching, but more as a practical tool to provide the necessary approximate DSM to support full-resolution, multi-image matching.

The paper commences with a short description of the NRP method as it could be applied in practice. An overview of the basic theory behind NRPs is then provided along with the description of the particular NRP selected by the authors for practical evaluation. This is followed by the discussion of image matching strategy for the final stage random pattern projection. In order to subject the NRP approach to experimental testing, a multi-sensor, on-line digital photogrammetric system is required and a description of the system developed for this purpose is given. Finally, results of practical testing of the NRP approach used for real-time featureless surface extraction are reviewed and the applicability of this method is discussed.

2 THE NRP APPROACH

2.1 The NRP Strategy for DSM Extraction

Shown in Figure 1 is a 4 x 16 NRP in which any point can be identified since the 2 x 3 binary sub-array to which it belongs appears only once. A more comprehensive account of NRPs will be provided in the following section. Here, it need only to be noted that should Figure 1 constitute an image, then each image point would be immediately recognisable, and each can be thought of as a unique coded target. Thus, if an object to be measured in a multi-camera digital photogrammetric network were to have such a pattern projected onto its surface, the XYZ co-ordinates of each point would be available through triangulation without the explicit need for image point correspondence determination.

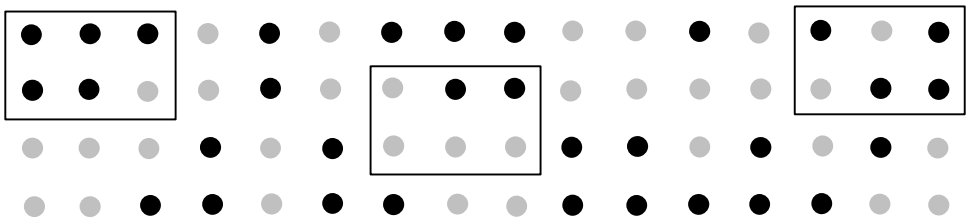

Figure 1: Projected 4x16 non-repeating pattern with uniquely occurring 2x3 sub-arrays.

As the point density of the NRP increases, so the resolution of the extracted DSM becomes finer. This allows for a new strategy of hierarchical DSM extraction, as illustrated in Figure 2. Shown in the figure are three levels of successively denser NRPs. The aim of each is to support fast, unambiguous object target detection and subsequent triangulation at progressively finer resolutions. It is not intended, however, that the full-resolution DSM would be extracted using an NRP. Instead, the texture projection coupled with traditional area- or feature-based matching would be employed at this final stage, the NRP-produced DSM offering the requisite preliminary surface model to support the matching.

From a practical standpoint, adoption of the NRP approach dispenses with the need for image pyramids and hierarchical image matching, since matching is performed only once using imagery of the projected texture pattern. As will be discussed, NRPs for practical application usually comprise only a modest number of distinct pattern codes. It is for this reason that a hierarchical projection system is needed in the NRP approach, since the second and subsequent projections may simply comprise multiple copies of the basic pattern. Ambiguity problems between adjacent NRP blocks will not arise in point correspondence determination due to geometric constraints imposed via both the DSM obtained at the previous projection stage and the exterior orientation of the multi-camera network. As an alternative to the hierarchical projection approach, a single NRP could well be incrementally moved over a given surface area by altering the projector orientation and thus providing a sufficiently dense DSM to support the final image matching stage.

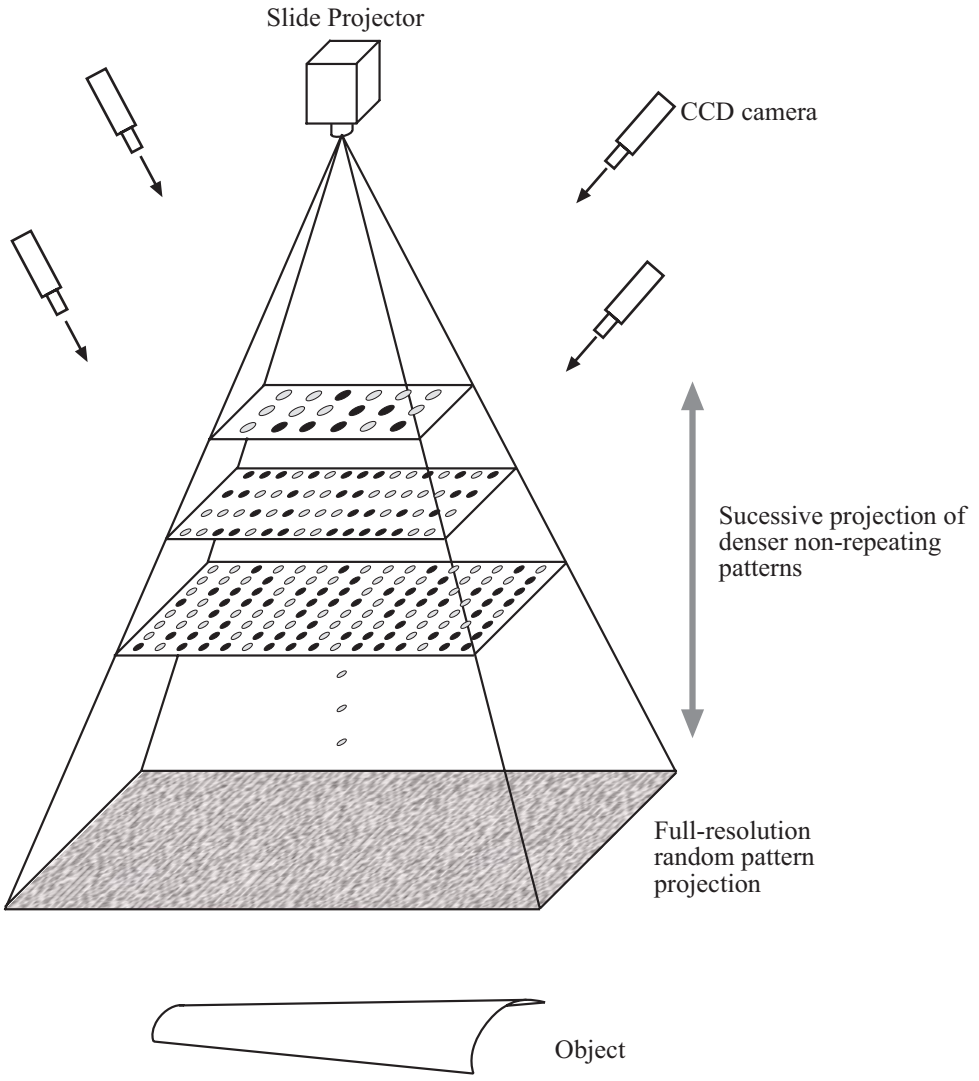

Figure 2: Hierarchical NRP strategy integrated into a 4-sensor vision metrology system.

2.2 Design and Recognition for NRPs

The principle of the NRP is closely related to the concept of Perfect Maps and Bruijn Sequences used in information theory and cryptology. A perfect map (*r, s; u, v)* is defined as a periodic *r* x *s* binary array in which every *u* x *v* binary array appears exactly once as a periodic sub-array (Paterson, 1994). In the case where r = 4, s = 16 and u = 2, v = 3, one of the possibilities for a perfect map is the target array shown in Figure 1, where the darker dots represent zero and the lighter dots one. Each 2 x 3 neighbourhood defines a code for the location of any target. The code is a six-element vector, which is unique for every target in the projected pattern. Once the targets have been recognised and their centres (or centroids) have been extracted, a data file of labeled 2D target centres can be generated. As unambiguous corresponding target co-ordinates from different images are obtained, the object point can be immediately triangulated.

An initial impetus for this investigation of the applicability of NRPs in photogrammetry was provided by the work of Nissanov et al. (1995) in which a multi-lettered 27 x 729 projection pattern, comprising uniquely occurring 3 x 3 sub-arrays, was employed in digital photogrammetric measurements of frozen rat brains. The essence of the NRP approach, as implemented here, is speed as opposed to accuracy. The actual targets comprising the projected NRP should therefore lend themselves to fast and easy detection and subsequent pattern recognition. Binary arrangements such as Figure 1 support both fast segmentation and extraction, as well as robust centroid determination to support triangulation. There are a host of different target designs that could be adopted, and for the reported investigation a pattern comprising projected dots of two distinct brightness levels has been employed.

The automatic procedure of extracting, measuring and labelling the points comprising the NRP image involves a number of stages. First, the image is scanned for the detection and extraction of seed points, which constitute candidate

targets. Second, for the bright and dim image 'blobs' comprising the NRP shown in Figure 1, a target validation process is carried out both to qualify targets through their geometric properties and to remove invalid seed points. The image co-ordinates of the centre of each valid blob are then determined via an intensity-weighted centroiding algorithm. At the fourth and final stage, the NRP is decoded and the individual points are labelled. Therefore, unambiguous image point correspondence are determined between the targets within the NRP for the different images forming the multi-image network. The entire operation should ideally occur within a second or so for each image. Given a known interior and exterior orientation, triangulation can generate 3D target coordinates to form the DSM which is used to support the final stage image matching.

3 IMAGE MATCHING STRATEGY

The final surface extraction is obtained via feature-based matching. In this operation, a random pattern projection is adopted to provide the texture. Both Lüyan (Lü, 1988) and Förstner operators (Förstner and Gülch, 1987) have been employed for interest point extraction. A technique based on monoplotting, and the epipolar geometry constraint were applied to aid in point correspondence determination.

3.1 Monoplotting Approach

The monoplotting strategy was first used for manual plotting of planimetric line maps from single photographs and the corresponding digital terrain model (DTM) data (Makarovic, 1973). In this research, it has been adopted as a tool for image matching. Under this approach, the object space position of each interest point in a 'reference' image is determined using the NRP derived DSM as indicated in Figure 3. Search windows for corresponding feature points in the remaining images are then determined by back projecting this object point into image space together with other consistency constraints.

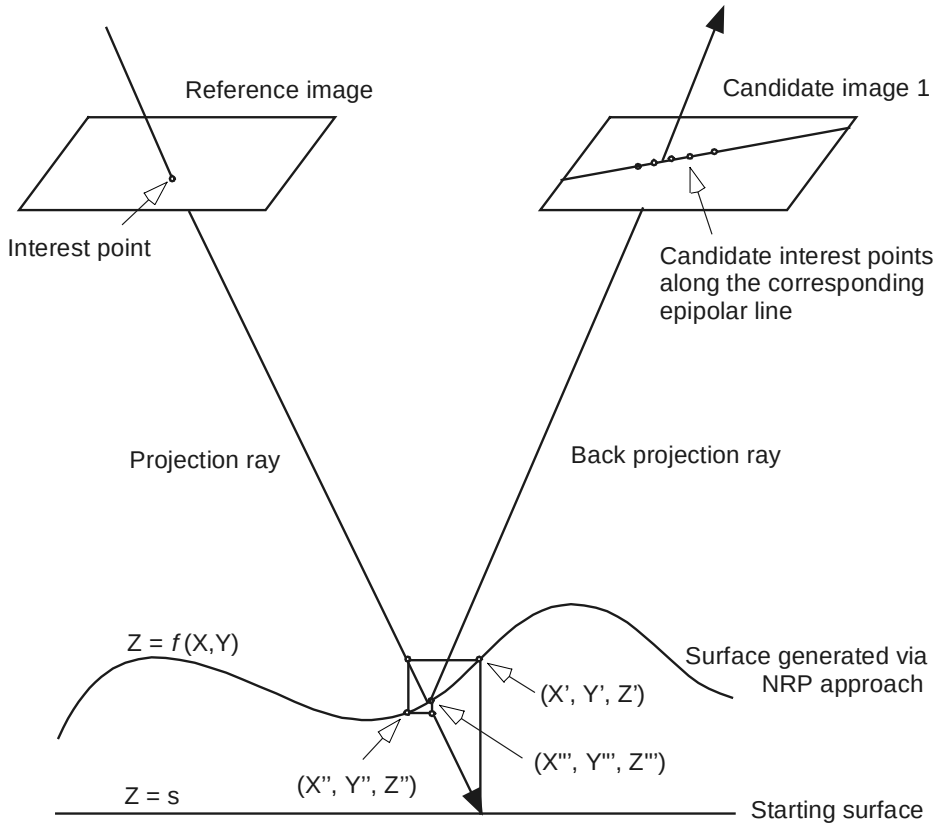

Figure 3: Monoplotting Strategy

Due to the sparseness of the DSM generated by the NRP approach, DSM interpolation is needed to locate the correct 3D coordinates for the projection rays from the reference image based on the selected reference data points of the DSM. The weighted arithmetic mean approach and the moving average method had been developed in the research.

The former algorithm has been adopted in the developed VMS as the dynamic point-wise interpolation method because of its robustness and computational economy.

Given the DSM generated by the NRP, and the developed interpolation approach, the monoplotting strategy (illustrated by Figure 3) can be implemented according to the following basic steps. These are premised on the assumption of known interior and exterior orientation (EO).

Step1: Project an interest point from the reference image to the object surface to determine (X′, Y′) given an arbitrary height Z = s, which is a plane close to the surface model (by estimation) used to initiate the iteration process. With respect to the collinearity equations, the relationship between the image point and the object point could be written as:

$$\begin{bmatrix} X - X_S \\ Y - Y_S \\ Z - Z_S \end{bmatrix} = \lambda^{-1}\mathbf{M}^{\mathrm{T}} \begin{bmatrix} x \\ y \\ -c \end{bmatrix} \tag{1}$$

where **M** is the rotation matrix and λ the scale factor. This expression can be expanded and transformed to:

$$\begin{aligned} \mathrm{X} &= \frac{m_{11}x + m_{21}y - m_{31}c}{m_{13}x + m_{23}y - m_{33}c} * (Z - Z_S) + X_S \\ \mathrm{Y} &= \frac{m_{12}x + m_{22}y - m_{32}c}{m_{13}x + m_{23}y - m_{33}c} * (Z - Z_S) + Y_S \end{aligned} \tag{2}$$

Which means, given height information Z, X and Y could be obtained.

Step2: Determine Z′=*f*(X′, ,Y′) via DSM interpolation.

Step3: Determine (X″, Y″) given Z′, using Eqs. 1-2 in Step1……

This iteration will not stop until the height difference ΔZ at the last two iterations (for instance, Z‴-Z″) is less then a threshold, or the process may cease when the number of iterations exceeds a given limit. At the end of the iteration process, a 3D DSM point is extracted, for example (X‴,Y‴,Z‴), which best describes the intersection of a projection ray from the reference image and the model Z = *f*(X,Y). In this DSM, the 3D point is back-projected to the candidate image. The intersection of this back-projection ray and the candidate image gives an approximate position of the homologous point. Obviously, the quality of this approximation is dependent on the density and the fidelity of the DSM generated by the NRP approach. It is also decided by the precision of the EO.

3.2 Epipolar Line Constraint

According to the principles of epipolar geometry, the homologous image point for any image must be located on its homologous epipolar line. Therefore the advantage of using epipolar geometry in image matching lies in the fact that the conventional two-dimensional correlation problem for homologous point searching will be substituted by a one-dimensional correlation search along the epipolar lines. The approximation generated by the back projection in the monoplotting provides a constraint for model consistency. The epipolar line constraint employed here gives further constraints and thresholding on the homologous point determination.

There are many ways to determine the corresponding epipolar lines in different images. These approaches can be based either on the coplanarity condition or on the principle of image projective transformation. Both of the approaches have been programmed for the image matching in this research. Epipolar line resampling has also been employed in this process.

3.3 Image Matching Implementation

The implementation of image matching between a reference and candidate image pair can be summerised as follows: Once provided with the NRP derived DSM, the developed software first performs a whole-image scan to extract the interest points in the reference image. It then projects interest points to the initial surface contour. The monoplotting operation presented in Section 3.1 extracts the intersection of the projection ray with the initial surface. The 3D coordinates of this point of intersection can then be projected back to the candidate image. Using the x-coordinate

information of this point in the candidate image, the corresponding epipolar line is obtained. A bundle of epipolar lines and a length threshold along the epipolar lines are both selected (the back projected point is taken as the centre of the threshold length). The one-dimensional correlation along the epipolar line is then conducted. The point-pair with the highest correlation coefficient is chosen as the matched-pair. Triangulation is finally used to extract the 3D information. Similar approaches are used in the image matching in the VMS four-camera network.

4 REAL-TIME VISION METROLOGY SYSTEM

In order to experimentally evaluate the NRP approach and the algorithms developed for image matching, it was necessary to integrate the surface measurement strategy into a real-time, multi-sensor VMS. Such a system had been developed within the Department of Geomatics at the University of Melbourne. This comprised four synchronised Pulnix TM-9701 CCD cameras, the images from which were transmitted via a multiplexer to a Bitflow Data Raptor framegrabber installed within a Pentium 200 PC. All necessary software for sensor calibration, exterior orientation and subsequent object point triangulation had been developed for this system, which was designed to operate as a 4-camera, real-time optical 3D co-ordinate measurement tool. Calibration in the developed VMS is afforded both via an off-line, multi-sensor self-calibration approach (Fraser *et al.* 1995), and by an in-place 'self-calibrating resection' which utilises a portable 3D control frame. In its standard configuration (e.g. Figure 2), the system could nominally accommodate objects of about 1.5m square and up to 1.5m in height.

The general functional model for sensor calibration and photogrammetric triangulation for the system can be written, in matrix form, as:

$$\mathbf{v} = \mathbf{A}_1 \mathbf{x}_1 + \mathbf{A}_2 \mathbf{x}_2 + \mathbf{A}_3 \mathbf{x}_3 - \mathbf{l} \tag{3}$$

where $\mathbf{v}$ is the vector of image coordinate residuals; $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ are vectors comprising the parameters for the exterior orientation, object space coordinates, and interior orientation elements and additional calibration parameters, respectively; $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ are the associated design matrices comprising the values of the partial derivatives of the collinearity equations with respect to the system parameters; and $\mathbf{l}$ is the vector of discrepancy values, namely the difference between the computed x, y image coordinates and their observed values.

The vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ and matrix $\mathbf{A}_3$ are detailed as follows:

$$\mathbf{x}_1 = [\ \Delta X_0 \quad \Delta Y_0 \quad \Delta Z_0 \quad \Delta\varphi \quad \Delta\omega \quad \Delta\kappa\ ]_i^T,\ i = 1, m \tag{4}$$

$$\mathbf{x}_2 = [\ \Delta X \quad \Delta Y \quad \Delta Z\ ]_j^T,\ j = 1, n \tag{5}$$

$$\mathbf{x}_3 = [\ x_p \quad y_p \quad dc \quad K_1 \quad K_2 \quad K_3 \quad P_1 \quad P_2 \quad b_1 \quad b_2]_k^T, \quad k = 1, nc \tag{6}$$

$$\mathbf{A}_3 = \begin{bmatrix} -1 & 0 & -\frac{\bar{x}}{c} & \bar{x}r^2 & \bar{x}r^4 & \bar{x}r^6 & \left(2\bar{x}^2 + r^2\right) & 2\bar{x}\bar{y} & \bar{x} & \bar{y} \\ 0 & -1 & -\frac{\bar{y}}{c} & \bar{y}r^2 & \bar{y}r^4 & \bar{y}r^6 & 2\bar{x}\bar{y} & \left(2\bar{y}^2 + r^2\right) & 0 & 0 \end{bmatrix}_{ijk} \tag{7}$$

Where m is the number of camera stations, n the number of points and nc the number of cameras, and $\bar{x} = x - x_p$ and $\bar{y} = y - y_p$.

To facilitate projection of both the NRPs and the final random texture to support image matching, a standard slide projector was employed, rather than a more sophisticated projection sequencing device which would be needed in a system tailored for industrial application. The different NRP slides were produced using the software Adobe Photoshop, and the texture slide was kindly provided by the INPHO Company of Stuttgart, Germany.

As regards accuracy, the 4-camera VMS configuration yields standard errors of object point triangulation of 0.15mm over an object distance of 2m (imaging scale of 1:220), for a standard error of image co-ordinate measurement of 0.1 pixel. It should be recalled, however, that accuracy is not a high priority for the coarse DSM determination associated with each NRP projection. What is sought is solely a surface model of sufficient resolution to support the final image matching operation.

5 EXPERIMENTAL TESTING

Following integration of the NRP approach and image matching strategies into the real-time vision metrology system described above, it was possible to experimentally evaluate these techniques for photogrammetric surface measurement. Two experimental objects were considered. The first, a 90 x 60 cm area of the car door shown in Figure 4, represented a smooth surface of low curvature. The second, a 40 x 70 cm surface area on the mannequin torso shown in Figure 5, constituted a surface of somewhat more complexity. Both objects were to be measured using three levels of NRP, followed by the final surface extraction using the projected texture pattern and feature-based matching.

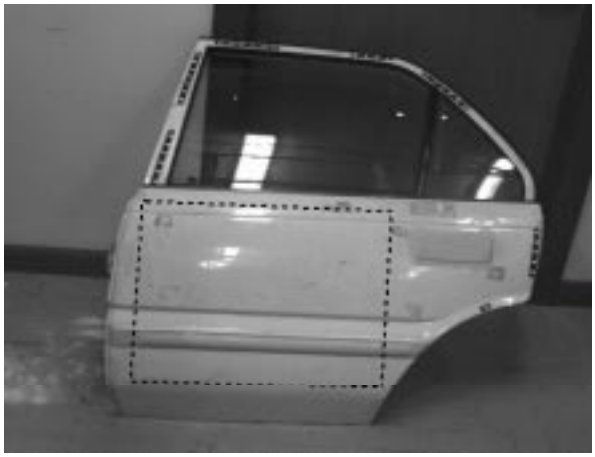

Figure 4: First test object, a car door, with the measured surface area indicated by the dashed line.

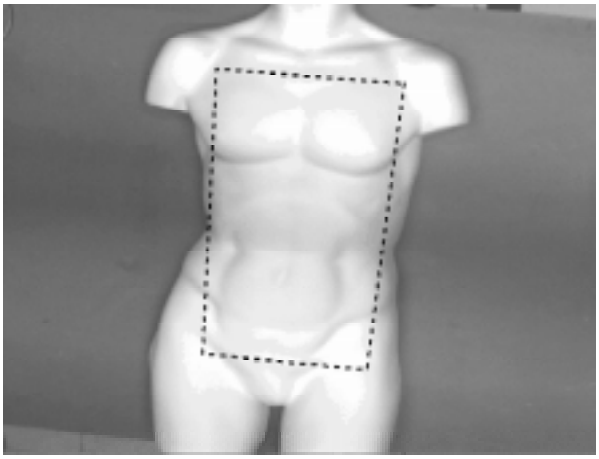

Figure 5: Second test object, a mannequin torso, with the measured surface area indicated by the dashed line.

The first NRP projection level consisted of a 4 x 4 pattern of bright and dim targets, with the second and third levels comprising patterns of order 8 and 12, respectively. The extracted DSM was then used to support the feature-based matching stage. Figure 6 and 7 provide a plot of the surface contours of the measured car door and mannequin torso. In each case the number of interest points utilised in the feature-based matching stage was close to 700. Indicated by crosses in Figure 8 is a 'height' profile through the final DSM produced for the car door by the NRP approach. Along with this is the profile (dots) measured by a commercial vision metrology system incorporating a touch probe, namely the V-STARS/M system from Geodetic Services, Inc. Although accuracy optimisation was not of primary concern in the experiments, it is noteworthy that in each case a representative standard error value of about 0.4 mm (approximately 0.3 pixels) was obtained in the surface measurements. Moreover, the measurement is a fully automatic procedure. No human intervention is needed in the course of operation.

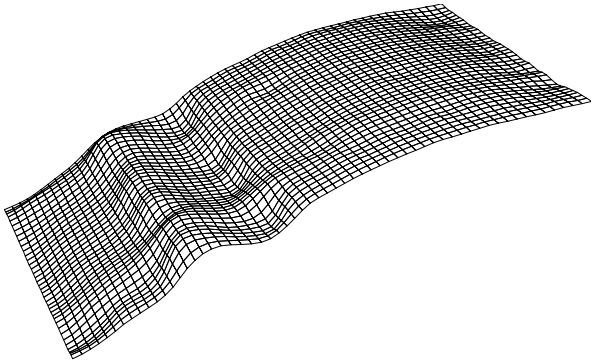

Figure 6: Surface model of the car door.

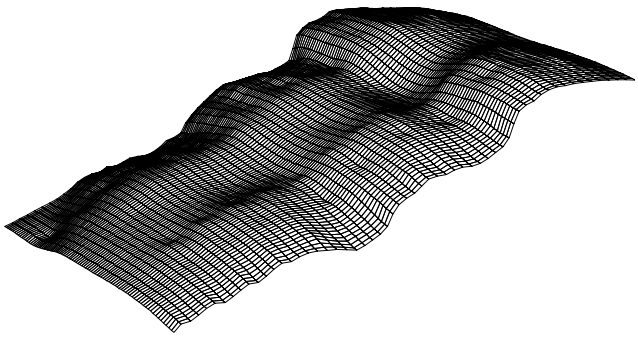

Figure 7: Surface model of the mannequin torso measured via the NRP approach

Although the aim of the NRP approach is to yield a fast and straightforward method for producing a preliminary DSM to support fine resolution image matching, it was not possible in the experimental testing to subject the timing aspect to careful evaluation. On the one hand, the NRP extraction software was not optimised in terms of computing time, and on the other the successive slide projections were carried out manually. The goal of developing a fully integrated and automatic system remained beyond the scope of the reported investigation, which focussed primarily on experimentally verifying the potential of the NRP approach. As a consequence, an operation that under ideal circumstances would yield a DSM in a matter of seconds instead consumed several minutes in the reported tests.

Although the NRP approach was successfully employed on the two test objects considered, limitations of the method became apparent during the course of the experiments. One limitation relates to the point density of the NRP. From the standpoint of measurement time, it is desirable to keep the number of projections to a minimum. In order to minimise the projection levels, however, more complex NRPs are needed and these will likely take additional time to decode. Also, occlusions can be problematic, as are surfaces with complex curvature or distinct changes in curvature (e.g. the raised trim on the car door).

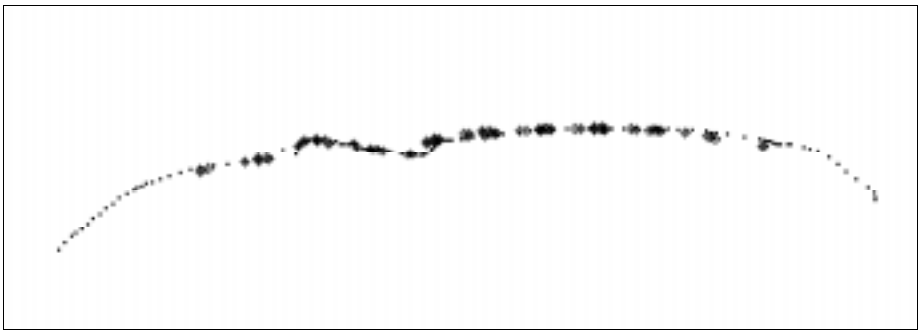

Figure 8: Comparison of the car door surface profiles extracted via the NRP strategy (crosses) and by a more precise vision metrology system (dots).

6 CONCLUDING REMARKS

The principal outcome of the two experimental measurements conducted was to demonstrate that the NRP approach and the feature-based image matching strategy could be integrated in a practical manner into a multi-sensor vision metrology system to provide an alternative approach to the photogrammetric measurement of featureless surfaces. In comparison to traditional image pyramid approaches to surface reconstruction via hierarchical multi-image matching, the NRP approach is a fully automatic, potentially faster and more robust means to derive a DSM of sufficient resolution to support a final precise image matching, be it feature- or area-based. The method is conceptually very straightforward and not difficult to implement, though multiple projections are required. The modest goals of the reported investigation were to develop the NRP approach, integrate it into an operational multi-sensor real-time vision metrology system, and demonstrate that practical application was feasible. These goals were by and large achieved.

REFERENCES

Ackermann, F. and Hahn, M., 1991. Image pyramids for digital photogrammetry. In: H. Ebner/D. Fritsch/C. Heipke (Eds.): Digital Photogrammetric Systems. Wichmann, Karlsruhe, pp. 43-58.

Förstner, W. and Gülch, E., 1987. A fast operator for detection and precision location of distinct points, corners and centers of circular features. Proceedings, Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken, Switzerland, 2-4 June, pp. 281-305.

Fraser, C.S., Shortis M.R. and Ganci G., 1995. Multi-sensor system self-calibration. In: S.F. El-Hakim (Ed.): *Videometrics IV*, Philadelphia, Pennsylvania, 25-26 October, *SPIE* 2598, pp. 2-18.

Grün, A.W. and Baltsavias, E.P., 1989. Automatic 3-D measurement of human faces with CCD cameras. Biostereometrics'88, Basel, SPIE 1030, pp. 106-116.

Jeschke, W., 1990. Digital close-range photogrammetry for surface measurement. In: A. Grün/E. Baltsavias (Eds.): Close-range Photogrammetry Meets Machine Vision, Zürich, Switzerland, 3-7 September, SPIE 1395, pp. 1058-1065.

Lü, Y., 1988. Interest-operator and fast implementation. International Archives of Photogrammetry and Remote Sensing, Kyoto, 27(3), pp. 491-500.

Makarovic, B., 1973, Digital mono-plotters, *ITC Journal*, 4, pp. 583-599.

Nissanov, J., Ozturk, C., and Kozinska, D., 1995. Brain surface scanning using structured light. In: S.F. Hakim (Ed.): *Videometrics IV*, Philadephia, 25-26 October, SPIE 2598, pp. 334-341.

Paterson, K.G., 1994. Perfect maps. IEEE Transactions on Information Theory, 40(3), pp. 743-753.

Trinder, J.C., Tjugiarto T., and Donnelly, B.E., 1990. A close-range digital photogrammetry system. In: A. Grün/E. Baltsavias (Eds.): Close-range Photogrammetry Meets Machine Vision, Zürich, Switzerland, 3-7 September, SPIE 1395, pp. 440-455.

Wong, K.W., Ke, Y., Slaughter, M., and Gretebeck, R., 1992. A computer vision system for mapping human bodies. International Archives of Photogrammetry and Remote Sensing, 29(5), pp. 850-855.