A Robust Close-Range Photogrammetric System for Industrial Metrology

Jian Xu*, Z.P. Fang*, Andrew A. Malcolm*, H. Wang**
*Machine Vision and Sensors Group, Gintic Institute of Manufacturing Technology,
71 Nanyang Drive, Singapore 638075
**EEE, Nanyang Technological University, Singapore

ABSTRACT

This paper describes a robust photogrammetric system for 3-D surface reconstruction. We use a linear camera calibration model with 14 camera parameters, which can be optimized using the least-squares solution. To address the non-linear lens distortion, a localized calibration concept is presented whereby different image locations have different sets of calibration parameters. A new subpixel target detection algorithm is proposed for accurate target center determination (with accuracy better than 0.02 pixel). The algorithm first uses zero-crossings of the second derivatives to determine an initial threshold for separating a target from its background. Then multiple threshold functions are selected to calculate multiple potential target centers. The weighted average of these potential centers results in a reliable final target position. Comparing with other subpixel algorithms, such as best-fit circle/ellipse, template matching based approach or single threshold gravity point method, our algorithm is more accurate and can be applied to any shapes of targets. To minimize the possibility of mismatch of targets among different views, we adopt a new reference point based matching approach. The target pattern from a projector contains a number of reference points. The stereo correspondences start from the reference points. The epipolar and position constraints are implemented to achieve error free matching, before the triangulation function is called to get the final 3-D coordinates. The system is able to achieve an accuracy (3σ) of 2.5μm (using calibration target with dimension 35mm x 35mm, 50mm lenses and Jai M1 CCD with 1300 x 1030 resolution).

1 INTRODUCTION

Over the past half decade, close-range photo-grammetry has become firmly established as a precise measuring tool for industrial metrology. Growing interest has been expressed in the need to obtain fast, accurate and reliable 3-D photo-grammetric measurements of scientific and engineering structures [1][2]. The fundamental principle used by photogrammetry is triangulation. By mathematically intersecting convergent lines in space, the precise 3-D location of a point can be determined. The involved metrology can be divided into five steps:

- Camera calibration to determine intrinsic and extrinsic parameters
- Apply targets to the object to be measured (either by using retro-reflective material or projector generated dot pattern)
- Taking photos from at least two views
- Target position detection
- Stereo matching and triangulation to get 3-D coordinates at target positions

In our previous paper [13] we used a calibration approach derived from a modified Haralick model [7]. By linearizing Haralick equations, we are able to describe the mapping between world coordinates and image coordinates with 14 parameters, which can be solved by the least-squares method. Experiments showed that our approach outperformed both DLT[5][6] and Tsai [3][4]models.

Precise subpixel target location is required in photogrammetry for stereo correspondence and triangulation. This can be accomplished by either center of gravity method, least squares template matching or edge based circle/ellipse fitting. Wong and Ho [8] developed the center of gravity formula based on the threshold window. However, target position determined by this formula is subjected to the variations in window size, position, and the value of the threshold. Later Eastman [9] improved the center calculation by introducing the weighted center of a target. The basic equations are

$$x = \frac{1}{M}\sum_{1}^{n}\sum_{1}^{m} j \times b(i, j) \quad (1a)$$

$$y = \frac{1}{M}\sum_{1}^{n}\sum_{1}^{m} i \times b(i, j) \quad (1b)$$

where

$$M = \sum_{1}^{n}\sum_{1}^{m} b(i, j)$$

and $b(i, j)$ is the binary image with value (0,1) at image position (i, j).

After the determination of target centers, traditional photogrammetry systems perform a "bundle adjustment" [10][11][12] to produce the final x,y,z coordinates of all measured target points. It involves the following steps

- Using coded targets for automatic target labeling (Fig. 1): the target design consists of a circle with two or more concentric rings containing 10 bit target id code
- A master calibration unit is imaged along with the object to be measured at different views
- Camera position and orientation are calculated from the master calibration unit images
- Triangulation for final 3-D coordinates

Fig. 1 Coded targets

In section 2.1, our calibration formula is derived from Haralick standard solution. The concept of the localized calibration is presented. In section 2.2, the mathematical description of the new subpixel target center detection is shown. Section 2.3 contains information on the proposed reference point based 3-D reconstruction. In section 3, we present some experimental results using our photogrammetry system. A comparison between different subpixel algorithms is made. Finally, we show the results of the reference point based 3-D reconstruction by measuring a 0.9m x 0.9m antenna surface. Section 4 presents conclusions regarding the system.

2 THE PHOTOGRAMMETRIC SYSTEM

The proposed photogrammetric system consists of the following three main components: camera calibration, subpixel target detection and reference point based matching and triangulation.

2.1 Camera Calibration

Our calibration scheme was described in detail in [13]. Here we briefly explain the basic concept. The calibration equations come from modified Haralick model [7]. The relationship between 3-D points in the world coordinate system (x_n, y_n, z_n) and their projections on the image plane (u_n, v_n) can be formulated as

$$a_1x_n + b_1y_n + c_1z_n + d_1x_nu_n + e_1y_nu_n + f_1z_nu_n + g_1x_nu_n^2 + h_1y_nu_n^2 + i_1z_nu_n^2 + j_1u_n + k_1u_n^2 + 1 = 0 \quad (2a)$$

$$a_2x_n + b_2y_n + c_2z_n + d_2x_nv_n + e_2y_nv_n + f_2z_nv_n + g_2x_nv_n^2 + h_2y_nv_n^2 + i_{21}z_nv_n^2 + j_2v_n + k_2v_n^2 + 1 = 0 \quad (2b)$$

Further simplification of equation (2) by omitting u_n^2, v_n^2 terms (since their coefficents are small), we have

$$a_1x_n + b_1y_n + c_1z_n + d_1u_n + e_1u_nx_n + f_1u_ny_n + g_1u_nz_n + 1 = 0 \quad (3a)$$

$$a_2x_n + b_2y_n + c_2z_n + d_2v_n + e_2v_nx_n + f_2v_ny_n + g_2v_nz_n + 1 = 0 \quad (3b)$$

Both Equations (2) (with 22 parameters) and (3) (14 parameters) are linear equations in term of camera parameters a_1, $a_2 \ldots$ d_1, $d_2 \ldots$., and can be solved using least squares method.

Localized calibration comes from the observation that the distortion of camera lenses is never symmetrical and thus cannot be represented by global parameters. Fig. 2 illustrates the idea of the localized calibration: by segmenting the total calibration area into, we are able to get the localized calibration parameters for each section which best describe the camera behavior there. We achieve this by feeding the equation (3) with world and image coordinates of calibration target points belonging that particular section. When reconstructing 3-D information after matching, localized calibration parameters will be used for triangulation.

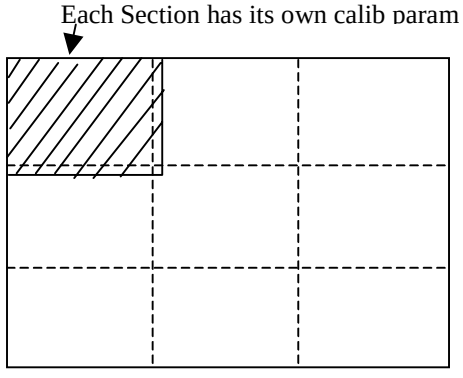

Fig.2 Example of localized calibration

2.2 Subpixel Target Center Detection

Fig. 3 shows a projector-generated target on an antenna surface with its 1-D cross section profile.

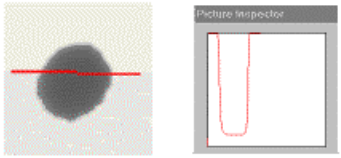

Fig. 3 A projector-generated target on an antenna surface with its 1-D cross section profile.

Let us examine the 1-D case first before extending the result to the 2-D situation. The objective is to determine a cutting line to separate the profile into target pixels and background pixels, as indicated in Fig. 4. We apply zero-crossings of the second derivative [14] to determine this cutting line. As illustrated in Fig. 5, the first derivative of the image function f(x) has a minimum and a maximum at the falling and rising edges. So the second derivative should produce two zero-crossings at the same positions.

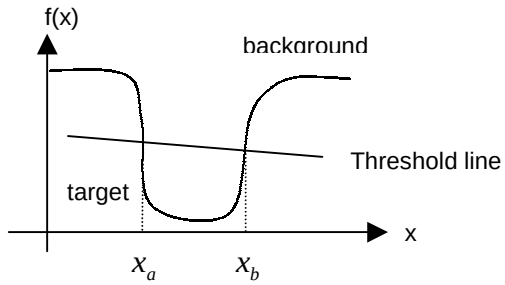

Fig. 4 A threshold line separates the profile into target and background.

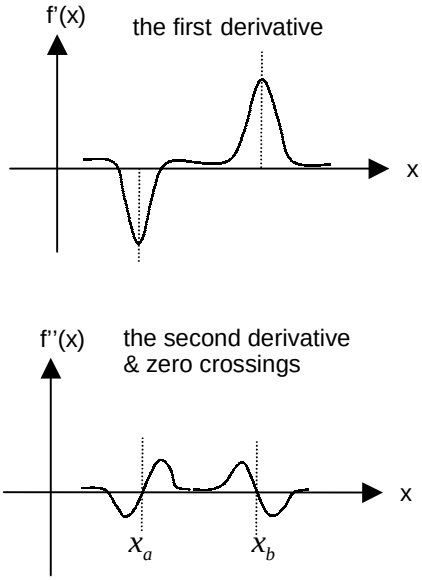

Fig. 5 Zero crossings at the edges

In this way the edge positions x_a, x_b and the corresponding $f(x_a), f(x_b)$ can be derived. The profile can be binarized as follow

$$b_0(x) = \begin{cases} 1 & f(x) < t_0(x) \\ 0 & f(x) >= t_0(x) \end{cases}$$

where threshold function

$$t_0(x) = \frac{(f(x_b) - f(x_a))(x - x_a)}{(x_b - x_a)} + f(x_a)$$

We get the initial guess of the target center:

$$x_{c0} = \frac{\sum_x x \times b_0(x)}{\sum_x b_0(x)}$$

Now instead of using a single cutting line, we can calculate a series of threshold functions $t_k(x)$ (see Fig. 6), binarization images $b_k(x)$ and target centers X_{ck} such that

$$t_k(x) = t_0(x) + k$$

$$b_k(x) = \begin{cases} 1 & f(x) < t_k(x) \\ 0 & f(x) >= t_k(x) \end{cases}$$

$$x_{ck} = \frac{\sum_x x \times b_k(x)}{\sum_x b_k(x)}$$

where $k = 0, \pm 1, \pm 2, \ldots \pm n$

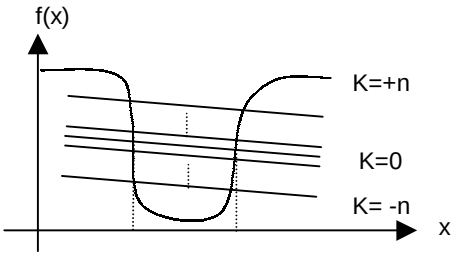

Fig. 6 Using multiple threshold lines to calculate multiple potential center points

Finally, the weighted average of all calculated x_{ck} is regarded as the target center

$$x_c = \frac{\sum_{k=0,\pm1..\pm n} w_k \times x_{ck}}{(2n+1) \times \sum_{k=0,\pm1..\pm n} w_k} \quad (4)$$

The weighting factors w_k are positive integers. We select w_k such that the position x_{c0} gets the highest factor 100. It decreases linearly to 50 at the positions $k = \pm n$.

The idea behind the selection of zero-crossings of the second derivatives as the best initial threshold is that this provides the widest dynamic range for as many cutting lines as possible.

The extension of equation (4) to the 2-D case is straightforward. The crucial question is how to compute the second derivative robustly in presence of noise. One possibility is to smooth a target image before applying the second derivative operation. This can be done by using a Gaussian filter

$$G(x,y) = e^{-\frac{x^2+y^2}{2\sigma^2}}$$

As pointed out in [14], the second derivative and the Gaussian filter operations can be combined and pre-calculated analytically. For example, the discrete approximation of a 5x5 second derivative of Gaussian filter can be expressed as

$$\nabla^2 G(x,y) = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ -1 & -2 & 16 & -2 & -1 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

After the convolution of the target image with $\nabla^2 G$, the locations in the convoluted image where the zero level is crossed correspond to the positions of edges. The grey values at all edge positions, together with their x,y positions are used to calculate a best fitting plane

$$t_0(x, y) = ax + by + c$$

Again, we can calculate a series of threshold functions $t_k(x,y)$, binarization images $b_k(x,y)$ and target centers x_{ck} and y_{ck}

$$t_k(x,y) = t_0(x,y) + k$$

$$b_k(x,y) = \begin{cases} 1 & f(x,y) < t_k(x,y) \\ 0 & f(x,y) >= t_k(x,y) \end{cases}$$

$$x_{ck} = \frac{\sum_x \sum_y x \times b_k(x,y)}{\sum_x \sum_y b_k(x,y)}$$

$$y_{ck} = \frac{\sum_x \sum_y y \times b_k(x,y)}{\sum_x \sum_y b_k(x,y)}$$

with $k = 0, \pm1, \pm2, \ldots \pm n$

Finally, the 2-D target center can be derived using

$$x_c = \frac{\sum_{k=0,\pm1..\pm n} w_k \times x_{ck}}{(2n+1) \times \sum_{k=0,\pm1..\pm n} w_k} \quad (5a)$$

$$y_c = \frac{\sum_{k=0,\pm1..\pm n} w_k \times y_{ck}}{(2n+1) \times \sum_{k=0,\pm1..\pm n} w_k} \quad (5b)$$

2.3 Reference Point Based Matching

The process of locating the same target point on an object in two different images is known as matching. Traditionally, photogrammetric systems use coded retro-reflective targets (Fig. 1) which can be decoded by the software. The proposed photogrammetric system uses a projector to generate targtes to achieve real-time processing.

A custom-designed algorithm based on reference points is introduced to achieve reliable matching results. Whilst it is difficult to blind search within two images for corresponding points, it will be easier to start from an already matched pair. To facilitate this, an unique, larger dot is projected at the centre of the object. This dot needs to be at least 50% bigger than the other dots in the projected array (see Fig. 8) . The matching is carried out by perfoming the following tasks (for simplicity, we assume only two views need to be matched):

1. Identify the reference points in the first and the second images
2. Match all reference points first if more than one reference points are defined (Using only epipolar and distance constraints only)
3. All targets are sorted according to their distance to the reference points

4. For all unmatched points in the first image, find the last matched point (at the beginning, this is the reference point). At this stage, the last matched point in the second image is also known
5. Calculate

   S_1 =Target size in the first image

   D_1 =Target distance between this unmatched point and the last matched point in the first image
6. For all points near the last matched point in the second image, calculate S_2, D_2 similarly and

   E_{12} =The distance of the target point in the second image to the epipolar line
7. Match is found if the following conditions are met:

   Distance constraint: $|D_1 - D_2| < T_d$ and

   Size constraint: $|S_1 - S_2| < T_s$ and

   Epipolar constraint: $E_{12} < T_e$,

   where T_d, T_s, T_e are thresholds
8. Repeat step 4 until all points are processed

3 EXPERIMENTAL RESULTS

Jai M1 CCD cameras with resolution 1300 x 1030 were used for our experiment.

3.1 Subpixel Algorithm

In table 1, we compare the results achieved by the proposed multiple threshold lines target center detection vs. other methods. The template matching based approach seems to be able to get nearly the same precision as our method. Nevertheless, it cannot handle the distorted targets. The performance of the proposed method has also been proved by a repeatability test. The repeatability of 10 measurements by multiple threshold method is better than 0.8µm, while other single threshold methods (excluding the template matching method) can only achieve 1.65µm.

Test	Edge based best fit Circle	Template matching (Matrox function)	Single threshold gravity method	Multiple Threshold lines (n=70)
1	1.7043	0.9127	1.6507	0.8989
2	1.6408	0.8891	1.5837	0.8654
3	1.3955	0.7235	1.3358	0.7340
4	1.6320	0.8512	1.6324	0.8430
5	1.5872	0.8380	1.5702	0.8352
6	1.4538	0.8656	1.5581	0.8182
7	1.6673	0.8934	1.6199	0.8768
8	1.7921	1.0735	1.7619	0.9335
9	1.2763	0.9128	1.2226	0.7063
10	1.8022	0.9951	1.7210	0.8887
Avg	1.5952	0.8955	1.5656	0.8394

Table 1 Measured absolute height standard deviation σ (unit µm) with best fitting circle approach, template matching based method, single threshold gravity method, and the proposed multiple threshold lines method. The calibration range is from 0-1000µm. σ is calculated based on 696 target points.

3.2 Reference Point Based Matching

In order to assess the practical accuracy of the system, it was proposed to measure the surface of a parabolic antenna dish approximately 0.9m in diameter. This required the construction and use of a 1m x 1m calibration target (Fig 7)

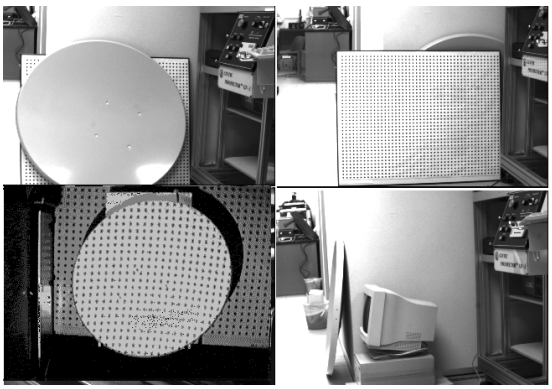

Fig 7: Top left: Antenna; Top right: Calibration unit; Bottom left: projected dot pattern;

A laser printer was used to print a calibration array of 32 by 26 circular dots. The accuracy of the pitch measurement was estimated at 0.4mm. The depth of the antenna dish was known to be approximately 250mm and as such the system was calibrated for an object located in the range 4500mm to 5000mm from the camera.

Fig. 8 shows the result of the reference point based matching. Fig. 9 shows the 3-D reconstruction.

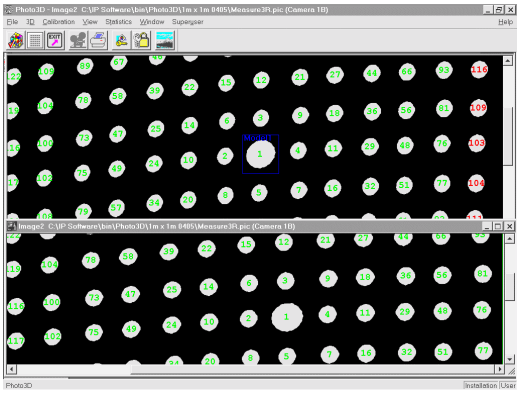

Fig. 8: The result of the reference point based matching: Green: successfully matched point; Red: no correspondence found.

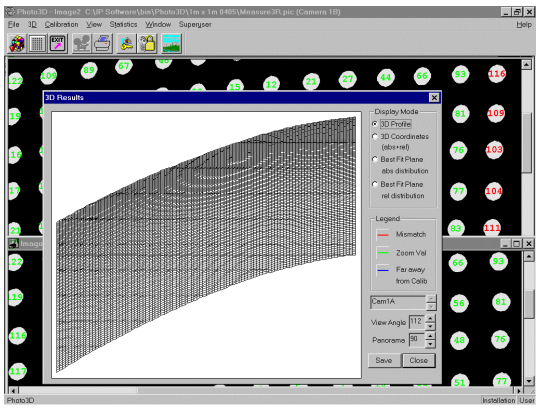

Fig. 9: Surface reconstruction

The accuracy of antenna surface measurement is estimably 0.5mm, due to the fact that the accuracy of 1m x 1m calibration unit is around 0.4 mm.

4 CONCLUSIONS

In this paper, we described a photogrammetric system using a projector to generate target points. Haralick camera calibration method was modified to enable a direct solution of all camera parameters with least-squares method. The localized calibration parameter sets were used to address the non-linearity at different lens positions. We introduced a new method to calculate target center reliably. Multiple thresholds were selected to calculate multiple potential target centers. The weighted averaging of these potential centers resulted in a reliable final target position. Comparing with other subpixel algorithms, such as best-fit circle/ellipse, template matching based approach or single threshold gravity point method, our algorithm is more accurate and can be applied to any shapes of targets. Furthermore, the reference point based matching made it possible to achieve error free labeling of target points in different views. Experiments showed that our system is robust and precise in 3-D reconstruction.

5 REFERENCES

[1] H. Haggrén, “Video image based photo-grammetry and applications”, SPIE Vol. 2350 Videometrics III (1994), pp 22-30.

[2] V. Kilpi, K. Sundelin, K. Halinen and E. Leikas, “Grinding of ship propellers with industrial robot”, Proceedings of the 25th International Symposium on Industrial Robots”, pp. 615-620, Hanover, 1994

[3] R. Tsai, "An efficient and accurate camera calibration technique for 3D machine vision", Proc. Conf. Computer Vision and Pattern Recognition, pages 364-374.

[4] R. Tsai and R. Lenz, "Techniques for calibration of scale factor and image centre of high accuracy 3D machine vision metrology", IEEE Trans. Pattern Analysis and Machine Intell., 10(5): 713-720, 1988.

[5] Y. I. Abdel-Aziz and H.M. Karara, “Direct Linear Transformation into object space coordinates in Close-Range Photogrammetry” in Proc. Symp. Close-Range Photogrammetry, Univ. of Illinois at Urbana-Champaign, Urbana, 1971, pp 1-18

[6] --, Photogrammetric potential of non-metric cameras, Civil Engineering Studies, Phtogrammetry Series 36, Urbana IL, Univ. of Illinois, 1974

[7] R. M. Haralick and L. G. Shapiro, “Computer and Robot Vision”, Vol. 2, Addison-Wesley, 1993

[8] K.W. Wong and W. Ho, “Close-range mapping with a solid state camera”, Photogrammetric Engineering & Remote Sensing, 52(1) 1986, pp. 67-74

[9] J.R. Eastman, “IDRISI for windows technical reference”, Graduate School of Geography, Clark University, Worcester, Massachusetts.

[10] D. Cosandier and M.A. Chapman, “High precision target location for industrial methology”, Videometrics I 1992, SPIE.

[11] C.S. Fraser, “Innovations in automation for vision metrology systems”, Photogrammetric Record, 1997, 15(90), pp.901-911

[12] Geodetic Services, Inc., V_STARS/INCA, [http://www.geodetic.com]

[13] J. Xu, Z.P. Fang, A. Malcolm and H. Wang, “Camera calibration for 3-D measurement with micron level accuracy”, Proc of 5th International Conference on Mechatronics Technology, pp. 56-61, 6-8 June, Singapore, 2001.

[14] M. Sonika, “Image Processing, Analysis, and Machine Vision”, PWS Publishing, 1999, pp. 83-87