

DIGITAL CLOSE RANGE PHOTOGRAMMETRY DATA ACQUISITION SYSTEM WITH COMBINED GPS GEOREFERENCE AND ITS CALIBRATION

Esmond Mok, Bruce King, Qing Zhu, Zhilin Li, Kent Lam

Department of Land Surveying and Geo-Informatics
The Hong Kong Polytechnic University, HK

Commission V, Working Group 2

KEYWORDS: CCD, Close Range Photogrammetry, GPS, Self-Calibration

ABSTRACT

A low cost digital close range photogrammetric system is presented for the acquisition of 3D surface information for efficient large scale and civil engineering mapping applications. Two GPS antennas are mounted together with two CCD cameras respectively on a stable bar and are employed for direct determination of camera position and azimuth during imaging process. This paper discusses the system calibration through a test field of control points. Some key issues about the integrated system such as the relationship between the GPS antennas and cameras and the achievable accuracy of the system are examined. The use of GPS geo-referencing of the cameras, multi-station convergent imaging and simultaneous adjustment are stressed for improving the accuracy and reliability of 3D surface data acquisition.

1 INTRODUCTION

Since small-format CCD cameras first attracted photogrammetrists' attention in the mid eighties these cameras have evolved high-resolution, large-area CCD cameras with digital output. The increase in resolution has resulted in a demand for greater accuracy and precision requirements for geo-referencing the camera system. In order to reduce the number of necessary control targets, position and attitude can be determined by integrating other information such as GPS and INS data and applying functional constraints. Obtaining complete, accurate, and fast description of a 3-D object or environment may not be possible without such integration.

GPS/INS systems has been successfully adopted in aerial photogrammetry for direct measurement of the exterior orientation parameters to decrease the amount of ground controls (Schwarz, 1995). Similarly, more sophisticated systems, and the methods for their integration have been developed for digital close range imaging (Li, 1997, Schwarz, 1996). Such research focuses on the areas of industrial and urban mapping applications. The system presented here is intended for the mapping of engineering sites, since civil engineering construction is a major activity in Hong Kong, and at present conventional mapping techniques are largely adopted despite its intrinsic inconvenience and danger on a construction site. It integrates the positioning technologies of GPS, photogrammetry and image processing.

Using such an integrated system, the basic problem is just to determine the relationship between GPS and camera. This process is referred to as the system calibration. The adopted approach is that of self-calibration as reviewed by Fraser (1997)

After briefly introducing the system configuration, a simple self-calibration model and DLT algorithm (Karara, 1989) are quoted directly. In the experiment analysis section, single-stereopair and multi-image processing and the imaging configuration are addressed.

2 SYSTEM CONFIGURATION

2.1 Cameras

Two Apple QuickTake 100 digital cameras are employed. These cameras take 8 hi-resolution (640 by 480 pixels) 24-bit color images with a pixel size of $8.75 \mu\text{m}$ by $8.75 \mu\text{m}$. Focus range is from 1.3m to infinity. The principal distance is approximately 8mm.

2.2 GPS system

Two Trimble 4600LS GPS receivers are used for camera positioning and one Trimble 4000SSI GPS receiver is used at base station for relative GPS positioning.

This configuration of Cameras/tripods/GPS antennas mounted on a common stable bar is illustrated as in Figure 1.

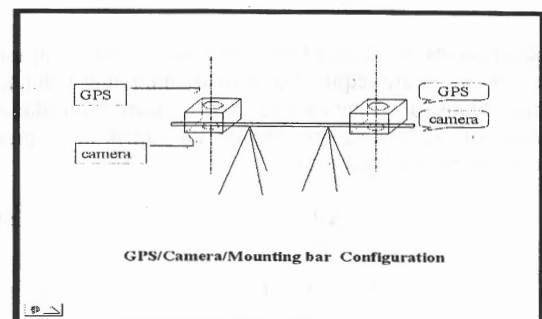


Figure 1 System configuration

3 SYSTEM CALIBRATION

Because it is attempted to directly use the GPS and calibrated results for camera orientation and location in the future work, for convenience the self-calibration method is adopted in our experiment. The following issues need to be considered:

1. The design, location and observation of the controls;

- The design and obtaining of the digital stereo pairs and GPS positions of cameras;
- The recognition and measurement of the targets images;
- The mathematical model of camera calibration that used is the self-calibrating bundle adjustment. After Fraser (1997) only the k_1 term is used for radial lens distortion.

$$\left. \begin{aligned} x-x_0+\Delta x &= -f \frac{a_1(X-X_s)+b_1(Y-Y_s)+c_1(Z-Z_s)}{a_3(X-X_s)+b_3(Y-Y_s)+c_3(Z-Z_s)} \\ y-y_0+\Delta y &= -f \frac{a_2(X-X_s)+b_2(Y-Y_s)+c_2(Z-Z_s)}{a_3(X-X_s)+b_3(Y-Y_s)+c_3(Z-Z_s)} \end{aligned} \right\} (1)$$

In Equation (1), (x_0, y_0) and f respectively are the principal point offsets and the principal distance of camera, $(\Delta x, \Delta y)$ is the image coordinate correction due to lens distortion, (x, y) and (X, Y, Z) are the coordinates of image point and its corresponding object point respectively, (X_s, Y_s, Z_s) are the coordinates of the exposure center, and the $(a_i, b_i, c_i) i=1,3$ are respectively the functions of three rotation angles (ω, ϕ, κ) :

$$\left. \begin{aligned} a_1 &= \cos\phi \cos\kappa \\ b_1 &= \cos\kappa \sin\phi \sin\omega + \cos\omega \sin\kappa \\ c_1 &= -\sin\phi \cos\omega \cos\kappa + \sin\kappa \sin\omega \\ a_2 &= -\cos\phi \sin\kappa \\ b_2 &= -\sin\kappa \sin\phi \sin\omega + \cos\omega \cos\kappa \\ c_2 &= \sin\kappa \sin\phi \cos\omega + \sin\omega \cos\kappa \\ a_3 &= \sin\phi \\ b_3 &= -\cos\phi \sin\omega \\ c_3 &= \cos\phi \cos\omega \end{aligned} \right\} (2)$$

$$\left. \begin{aligned} \Delta x &= k_1 \times x_r r^2 + k_2 \times x_r r^4 + p_1 \times (2x_r^2 + r^2) + p_2 \times 2x_r y_r \\ \Delta y &= k_1 \times y_r r^2 + k_2 \times y_r r^4 + p_1 \times 2x_r y_r + p_2 \times (2y_r^2 + r^2) \\ x_r &= x - x_0 \\ y_r &= y - y_0 \\ r^2 &= x_r^2 + y_r^2 \end{aligned} \right\} (3)$$

These equations result in 10 unknown parameters. That means at least 6 targets are required to provide minimum redundancy during calibration of one camera. After linearisation, the error equations of the least squares bundle adjustment are expressed as following matrix equations:

$$V = AX - L \quad (4)$$

$$\left. \begin{aligned} L &= (x-x' \quad y-y')^T \\ X &= (X_s \quad Y_s \quad Z_s \quad \phi \quad \omega \quad \kappa \quad f \quad x_0 \quad y_0 \quad k_1)^T \\ A &= \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} & a_{18} & a_{19} & a_{110} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & a_{27} & a_{28} & a_{29} & a_{210} \end{bmatrix} \end{aligned} \right\} (5)$$

In Equation (5), x' and y' are the image coordinates computed by means of approximate values of the parameters according to the collinearity equations; a_{ij} ($i=1,2, j=1,10$) are the partial derivatives with respect to the 10 unknowns for x and y . Based on the least square adjustment, the normal equations and their solution can be obtained as Equation (6).

Since the convergence of the least squares solution depends on the initial values of exterior parameters, a DLT algorithm (Karara, 1989) is used to provide this information. The DLT equations with 11 unknowns from which all the camera parameters can be derived are expressed as Equation (7).

$$\left. \begin{aligned} A^T A X &= A^T L \\ X &= (A^T A)^{-1} A^T L \end{aligned} \right\} (6)$$

$$\left. \begin{aligned} x - \delta_x &= \frac{L_1 X + L_2 Y + L_3 Z + L_4}{L_9 X + L_{10} Y + L_{11} Z + 1} \\ y - \delta_y &= \frac{L_5 X + L_6 Y + L_7 Z + L_8}{L_9 X + L_{10} Y + L_{11} Z + 1} \end{aligned} \right\} (7)$$

$$L = -1 / \sqrt{(L_9^2 + L_{10}^2 + L_{11}^2)} \quad (8)$$

$$\left. \begin{aligned} x_0 &= (L_1 \times L_9 + L_2 \times L_{10} + L_3 \times L_{11}) L^2 \\ y_0 &= (L_5 \times L_9 + L_6 \times L_{10} + L_7 \times L_{11}) L^2 \\ f_x &= \sqrt{[(L_1^2 + L_2^2 + L_3^2) L^2 - x_0^2]} \\ f_y &= \sqrt{[(L_5^2 + L_6^2 + L_7^2) L^2 - y_0^2]} \end{aligned} \right\} (9)$$

$$\left. \begin{aligned} f &= (f_x + f_y) / 2 \\ \omega &= \tan^{-1} (-L_{10} / L_{11}) \\ \phi &= \sin^{-1} (L_9 \times L) \\ a_1 &= L(x_0 \times L_9 - L_1) / f_x \\ \kappa &= \cos^{-1} (a_1 / \cos\phi) \end{aligned} \right\} (10)$$

$$\begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix} = \begin{bmatrix} L_1 & L_2 & L_3 \\ L_5 & L_6 & L_7 \\ L_9 & L_{10} & L_{11} \end{bmatrix}^{-1} \begin{bmatrix} L_4 \\ L_8 \\ 1 \end{bmatrix}$$

- The relationship between cameras and GPS antennas. Following calibration the distances and angles among cameras and GPS antennas can be considered as fixed values, They can then be used to reconstruct the imaging geometry for subsequent photogrammetric intersection.

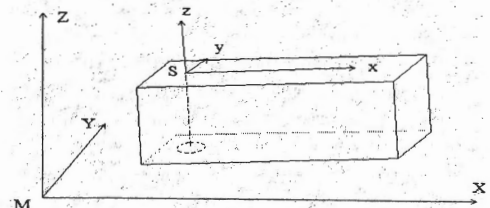


Figure 2 the relationship between mapping frame and sensor frame

As illustrated in Figure 2, the relationship between mapping frame (M-XYZ) and the sensor frame (S-xyz) can be expressed by six parameters, i.e. 3 rotations and 3 positions, because the

whole sensor system can be considered as a rigid body. The GPS observations can yield part of these parameters, and the differential coordinates between the GPS antennas and the cameras then can be obtained based on the calibration results (if the sensor system is leveled). In the future mapping process, the camera positions are able to compute just based on the GPS observations and these results

If no rotation about the x-axis, then there are only two rotations between sensor frame and mapping frame, i.e. φ (clockwise about y-axis) and κ (clockwise about z-axis). The rotation matrix can be expressed as Equation (11).

$$R = \begin{pmatrix} \cos \varphi \cos \kappa & \sin \kappa & -\sin \varphi \cos \kappa \\ -\cos \varphi \sin \kappa & \cos \kappa & \sin \varphi \sin \kappa \\ \sin \varphi & 0 & \cos \varphi \end{pmatrix} \quad (11)$$

Let the left GPS antenna position as the origin of sensor frame and its coordinates in the mapping frame as (X_0, Y_0, Z_0) , the coordinate of any position such as the camera center in sensor frame as $(\Delta X, \Delta Y, \Delta Z)$. The purpose of calibration is just to determine these values. If the calibrated coordinates of sensor position in mapping frame is (X, Y, Z) , then the coordinates of any position in sensor frame can be computed as following expressions:

$$\begin{pmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{pmatrix} = R^{-1} \begin{pmatrix} X - X_0 \\ Y - Y_0 \\ Z - Z_0 \end{pmatrix} \quad (12)$$

In the future work, by means of the two GPS observations, the rotation matrix R can be determined, and the camera position then can be computed according to above calibrated values. As the same, the differential rotation between sensor frame and camera exterior orientation also can be obtained by the calibration. Thus, based on this differential rotation and the differential rotation between S-xyz and M-XYZ, the exterior orientation of camera can be determined at exposure time without field control.

6. Forward intersection of any object point corresponding to their stereo images. This is achieved as follows:
 - Compute the rotated 3 dimensional coordinates of image points based on the solved camera parameters:

$$\left. \begin{aligned} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_r &= \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}_r \times \begin{pmatrix} x \\ y \\ -f \end{pmatrix}_r \\ \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}_l &= \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}_l \times \begin{pmatrix} x \\ y \\ -f \end{pmatrix}_l \end{aligned} \right\} \quad (13)$$

- As the inversion of collinearity equations, including unknown scale factors K1 and K2 for each point from the two conjugate images, giving:

$$\left. \begin{aligned} X &= K_1 \times X_l + X_s = K_2 \times X_r + X_s \\ Y &= K_1 \times Y_l + Y_s = K_2 \times Y_r + Y_s \\ Z &= K_1 \times Z_l + Z_s = K_2 \times Z_r + Z_s \end{aligned} \right\} \quad (14)$$

This yields six equations for 5 unknown parameters. The least squares solution yields the object space coordinates of the required point.

4 EXPERIMENT

4.1 Test Data Description

A three dimensional test range of 23 retro-reflective targets was established on an A-shaped canopy, associating pipe and background wall. The target array occupies a volume of approximately 5m by 4m by 4m (depth). The test range is shown in Figure 3.



Figure 3 the picture of test range

Three stations were positioned by the Trimble GPS system to a precision of 1mm. these stations form the control points from which all the targets were intersected by means of one second theodolite (Wild T2). The estimated precision of the target points was ± 5 mm.

The GPS observations for camera location (which were gathered over a period of 1 hour for each stereopair by relative GPS method) were transformed from WGS-84 Cartesian coordinates to Hong Kong 1980 local coordinates.

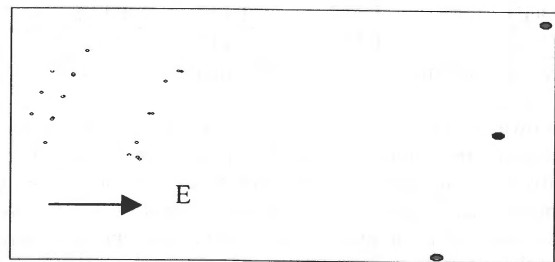


Figure 4 Illustration of the control range

The accuracy and precision of the geo-referencing process was evaluated by an experiment. Three stereopairs of the test field were taken. Of the 23 targets, 12 serve as control information for the calibration, with the others being used as check points. The test images were taken from 10m to 15m distance away with the base line of about 1.1m. Moreover, as illustrated in Fig.4, the N-axis of the object space coordinate system is approximately parallel to the camera base. On the other word, the light ray is nearly parallel to the E-axis. The precision of image coordinate measurement was about 0.5 pixel.

The resulting data was processed both as single stereopair and as multiple stereopairs to examine the potential accuracy and precision improvement.

4.2 Single Stereopair Processing

The camera parameters resulting from the self-calibration are shown in Table 1. As the image units used are in pixels and considering the pixel size is $8.75 \mu m$, then it's easy to see that the calibrated camera interior parameters are very closed to those mentioned in section 2.1.

Table 1 Camera Parameters

Para.	Self-calibration	
	camera 1	camera 2
fx	916.885	920.403
fy	916.885	920.403
x0	326.939	325.119
y0	247.464	239.608
K1	1.829e-7	1.959e-7

The relationship between the camera center and GPS antenna can be determined through distances as shown in Table 2. Only distance relationships are constant as the bar is moved, rotated or tilted during the image acquisition. For the purposes of this experiment, the camera/GPS mounting bar is considered as a rigid body (even though because the reason of material used, the bar was still distorted during photography). For further dynamic mapping purposes, additional sensors for tilt orientation of the bar maybe needed.

In this experiment, the distance relationship between the cameras and the GPS antennas in fact are preserved, but the calibration results and the GPS surveyed results of the three stereopairs are quite different from each other. This is just a result of the image coordinate measurement, GPS observations and the limited control accuracy.

Table 2 Relationship between GPS and camera (m)

parameter	station1	station2	station3
DGleftCleft	0.139	0.122	0.139
DGrightCright	0.093	0.152	0.160
DGleftCright	1.135	1.161	1.165
DGrightCleft	1.125	1.108	1.075
DGG	1.117	1.137	1.118
DCC	1.138	1.117	1.104

where, D = distance, G = GPS antenna, C = camera

As shown in Table 3, the effect of radial lens distortion is significant. Inclusion of the K1 parameter results in an improvement of between 30% and 50% over not using the parameter. That means that the geometric distortion of current CCD camera is an important factor what must be considered. About the effect of higher order terms of lens distortion will be investigated in the coming experiment. In this experiment the accuracy of control point coordinates can reach only 5mm and the accuracy of image coordinates measurement is about half pixel. Due to the limitation of the camera resolution and the blurredness effect, even though imaging in the evening to enhance the background/foreground contrast, the accuracy improvement of image coordinate measurement is still very insignificant.

Table3 The accuracy assessment of calibration (m)

	DLT	Self-calibration			
		#	k1	k1+k2	k1+k2+p1+p2
N	0.014	0.014	0.007	0.008	0.008
E	0.072	0.063	0.054	0.053	0.055
H	0.012	0.010	0.009	0.007	0.008

The RMSE is computed based on the check points which are not included in the calibration process.

4.3 Multi-image Simultaneously Processing

Since the imaging baseline length is limited by the mounting bar, it is hard to improve the accuracy of spatial intersection depending on only one stereopair. The same three stereopairs are then processed simultaneously. The resulting relationships between GPS and Camera are shown in Table 4.

Table 4 Relationships from multistation adjustment (m)

parameter	station1	station2	station3
DGleftCleft	0.1215	0.1222	0.1205
DGrightCright	0.1720	0.1588	0.1784
DGleftCright	1.1338	1.1689	1.1560
DGrightCleft	1.1269	1.1088	1.0874
DGG	1.1169	1.1365	1.1182
DCC	1.1272	1.1253	1.1072

Comparing these results with those of Table 2, the RMSE (computed from the difference of the offset and its average value from three stereopairs) of the latter is 0.0108m more less than that of 0.0175m. Also the accuracy of spatial intersection is improved evidently as shown by the comparison of Table 3 and Table 5. For example the relative accuracy of position (N,H) comparing with the size of object is improved from about 1/310 to 1/700, and the relative accuracy of depth is also improved from 1/180 to 1/1100.

Table5 The accuracy of multistation adjustment (m)

RMSE	Case I	Case II
N	0.0092	0.0049
E	0.0250	0.0091
H	0.0034	0.0029

where, in the Case I the three camera stations are in a horizontal plane, and in another case they are in a vertical plane.

4.4 The Effect of Imaging Configuration on the Accuracy

Because the network configuration is also very important to the accuracy of intersection, in order to achieve higher mapping precision by multi-station imaging, careful attention must be paid to the design of the photogrammetric network (Fraser, 1996). In order to analyze the effect of configuration on the photogrammetric accuracy, there were two different station configurations designed. The first one including three stereopairs is illustrated in Figure 5, in which the height differences of these three stations were very small, i.e. they were set almost in a horizontal plane; another group of convergent images were taken as shown in Figure 6, in this case, the stations were set in a vertical plane. Since taking more robust imaging configuration, the accuracy is increased by about 40% to 60% as shown in Table 5. In such case of configuration, the relative depth accuracy can reach 1:1100.

From Table 6, it is also easy to see that the baseline/distance ratio is a very important factor of configuration to the photogrammetric data quality. To increase the baseline will cause to decrease the overlap, so in our experiment the largest convergent angle can be only about 42 degrees. On the other hand, the distance from object to camera is also the key factor on mapping accuracy. However because of the limited dimension of the test site on the building top, such experimental analysis with respect to the imaging distance has not been done.

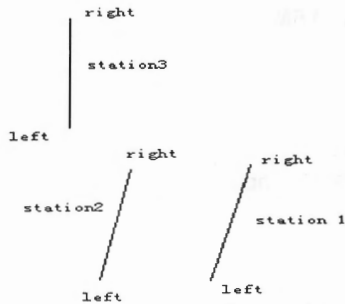


Figure 5 case I: horizontal spatial intersection

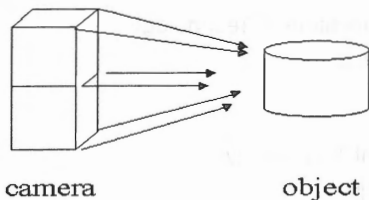


Figure 6 case II: solid spatial intersection

Table 6 The effect of baseline/convergence angle on the accuracy of single-stereopair adjustment (m)

Baseline	1.107	2.762	7.141
Convergence Angle	6	16	42
N	0.0101	0.0071	0.0061
E	0.0531	0.0252	0.0105
H	0.0081	0.0032	0.0031

5 CONCLUSIONS

Using the low cost system designed in our case, since the baseline/distance ratio is very small, the relative accuracy of single-stereopair processing is then only about 1/200 to 1/400. But if better baseline/distance ratio and multi-stations adjustment are concerned, this relative accuracy can be increased to 1/700 to 1/1000. Because of the using of GPS for geo-referencing purpose, no control targets will be needed in practice and multi-image with strong configuration can be obtained conveniently, and through multi-stations simultaneous bundle adjustment higher accuracy and reliability can be achieved. It even enables the real-time mapping of complicated construction site.

6 ACKNOWLEDGMENT

This research is supported by the Hong Kong Polytechnic University (351/616). The authors would like to thank Mr. Lawrence Lau for his assistance in GPS surveying and data processing.

7 REFERENCES

- Fraser, C.S., (1997). Digital camera self-calibration, *ISPRS Journal of Photogrammetry and Remote Sensing*, 52(1), pp.149-159.
- Fraser, C.S., (1996). Network Design, *Close Range Photogrammetry and Machine Vision*, Whittles Publishing, pp.256-281.
- Karara, H.M., (1989). *Non-Topographic Photogrammetry*, American Society for Photogrammetry and Remote Sensing.
- Li Rongxing, (1997). Mobile Mapping: An Emerging Technology for Spatial Data Acquisition, *Photogrammetry Engineering & Remote Sensing*, 63(9), pp.1085-1092.
- Schwarz, K.P., (1995). Integrated Airborne Navigation Systems for Photogrammetry, *The Wichmann Verlag, Photogrammetric week '95, Stuttgart*, pp.139-153.
- Schwarz, K.P., and N. El-Sheimy, (1996). Kinematic Multi-Sensor Systems for Close Range Digital Imaging, *International Archives of Photogrammetry and Remote Sensing*. XXXI(B3), pp.774-785.