System Structure and Calibration Models of Intelligent Photogrammetron

PAN Heping  ZHANG Chunsen

1 Introduction

In order to break through the limitations of the current dominant digital photogrammetric systems, photogrammetron has been proposed recently as a new class of intelligent photogrammetric systems. It is designed to be an active stereovision system driven by intelligent software agent architecture, aiming at realizing a number of newly defined functionalities of intelligent photogrammetry. Some main functionalities that go beyond the traditional photogrammetry and the currently dominant digital one include real-time photogrammetry in video surveillance, photogrammetry-enabled robots, intelligent multi-camera network for close-range photogrammetry. Photogrammetron I, the first type of photogrammetron, is designed to be a coherent stereo photogrammetric system in which two cameras are mounted on a physical base, similar to a head-eye system in robot vision, but the stereo camera baseline length is changeable.

The calibration of photogrammetron is far more complicated than just calibrating the cameras in traditional photogrammetry because photogrammetron posses a self-contained and auto-controlled physical structure driven by the intelligent agent software architecture.

Physically, photogrammetron I, as shown in Fig. 1, is made up of a physical support base called ‘shoulder’, a pan-tilt unit called ‘head’ mounted on the shoulder, a plate called ‘stereo camera plate’ or ‘stereo plate’ mounted on the head, the left and right cameras with their pan-tilt unit on top of the stereo base. Each pan-tilt
unit has two angular freedoms; pan and tilt. In total, there are nine freedoms; pan and tilt angles for each of the three pan-tilt units, the baseline length between two cameras, the focal length of each of the two cameras. Besides these freedoms, there are still a number of prefixed system parameters such as the geometry between the head and the stereo base, between the stereo base and each of the camera pan-tilt units and between each camera and its pan-tilt unit. Therefore, the whole parameter set of the system can be divided into two subsets; the free parameters and the fixed parameters.

The system calibration of Photogrammetron I is divided into two parts; the determination of the fixed parameters and that of the free ones. The calibration for the fixed parameters can be done in a laboratory in advance of the system operation, which is called the ‘in-lab’ calibration. The calibration for the free parameters has to be done in real-time during the system operation, which is called the ‘in-situ’ calibration.

2 A geometric model of Photogrammetron I

A basic structure of Photogrammetron I consists of five hardware parts; the shoulder, the head, the stereo plate, the left camera and its pan-tilt unit and the right camera with its pan-tilt unit. Each of them can be described in more details as follows.

2.1 The shoulder

The shoulder supports the system. It can be a still tripods or a vehicle with wheels or robot with legs. For the time being, we just assume that the shoulder stays still relative to the surveillance environment. For this part, a Euclidean reference system $O-XYZ$ is assumed, where the $Z$ axis corresponds to the vertical line pointing from the bottom to the top through the centre of the shoulder.

2.2 The head

The head is a pan-tilt unit mounted on the top of the shoulder. Relating to the shoulder, the head can pan an angular freedom $\phi$ around the $OZ$ axis. It can also tilt an angular freedom $\theta$, which is orthogonal to the pan angle $\omega$.

2.3 The stereo plate

The stereo plate is a plate to support the two stereo cameras. It is fixed on the top of the head. On the top of the stereo plate the left and right camera pan-tilt units are symmetrically mounted. For simplicity, we shall call the left/right pan-tilt unit supporting left/right camera and left/right unit. Since the stereo plate is fixed on the top of the head, it can tilt an angle $\phi$. A reference system $S-UVW$ is assumed for the stereo plate. The origin $S$ is on the $OZ$ axis and with a distance $h$ from the origin $O$. The $S-U$ axis is horizontal pointing from left to right, the $S-V$ axis refers to the depth from the system toward the objects, and the $S-W$ axis is pointing upwards. The transformation from $S-UVW$ to $O-XYZ$ is defined by

$$
\begin{bmatrix}
X \\
Y \\
Z
\end{bmatrix} =
\begin{bmatrix}
X_s \\
Y_s \\
Z_s
\end{bmatrix} +
R
\begin{bmatrix}
U \\
V \\
W
\end{bmatrix}
$$

where
2.4 The left camera and its pan-tilt unit

On top of the stereo plate, the left and right pan-tilt unit are placed along the S-U axis, and they are symmetrically placed about the centre, the S-W axis. Let C denote the perspective center of the left camera, and \( f \) the focal length. A reference system \( C-xyz \) is assumed for the left camera. The C-z axis is the principal axis of the camera pointing through the perspective center \( C \) towards the scene. The image plane is at the back of \( C \). An image point is positioned with coordinates \((x, y, -f)\). The principal point is located at \((X_C, Y_C, -f)\). For the left unit supporting the left camera, there are the geometric centre \( T \), the pan angle \( \alpha \) and the tilt angle \( \beta \). We must be aware that the perspective centre \( C \) and the unit centre do not coincide. And due to the discrepancy between the two centres, the perspective centre \( C \) is a function of the pan and tilt angles \( \alpha \) and \( \beta \) as well as the focal length \( f \), which may be expressed generally as

\[
C = C(T, \alpha, \beta, f) \tag{5}
\]

A simple form of this function in the stereo plate reference system \( S-UVW \) is

\[
\begin{bmatrix}
U_C \\
V_C \\
W_C
\end{bmatrix} =
\begin{bmatrix}
U_T \\
V_T \\
W_T
\end{bmatrix} + d
\begin{bmatrix}
R_x \cdot R_y \\
R_y
\end{bmatrix}
\tag{6}
\]

where \( a, b, c, d \) are constants and fixed once the camera is fixed on the unit, and \( R_x, R_y \) are two two-dimensional rotation matrices

\[
R_x = \begin{pmatrix}
\cos \alpha & -\sin \alpha & 0 \\
\sin \alpha & \cos \alpha & 0 \\
0 & 0 & 1
\end{pmatrix} \tag{7}
\]

\[
R_y = \begin{pmatrix}
1 & 0 & 0 \\
0 & \cos \beta & -\sin \beta \\
0 & \sin \beta & \cos \beta
\end{pmatrix} \tag{8}
\]

and

\[
R_z = \begin{pmatrix}
1 & 0 & 0 \\
0 & 0 & 1 \\
0 & 1 & 0
\end{pmatrix} \tag{9}
\]

Note that the image coordinate system generally has a rotation about the principal axis, which is denoted here by \( \gamma \). The transformation from the image coordinates to the stereo plate reference system is defined by

\[
\begin{bmatrix}
U \\
V \\
W
\end{bmatrix} =
\begin{bmatrix}
R_x \cdot R_y \\
R_y
\end{bmatrix}
\begin{bmatrix}
U_C \\
V_C \\
W_C
\end{bmatrix} +
\begin{bmatrix}
x - x_c \\
y - y_c \\
-f
\end{bmatrix} \tag{10}
\]

where

\[
R_y = \begin{pmatrix}
\cos \gamma & -\sin \gamma & 0 \\
\sin \gamma & \cos \gamma & 0 \\
0 & 0 & 1
\end{pmatrix} \tag{11}
\]

2.5 The right camera and its pan-tilt unit

Similarly we have everything for the right camera and its pan-tilt unit. Any element on the right camera or pan-tilt unit is denoted by \( x' \) corresponding to its counter part \( x \) on the left camera or unit. Therefore, for the right camera we have the perspective center \( C' \), the reference system \( C'-x'y'z' \), the focal length \( f' \) and the principal point \((x'_C, y'_C, -f')\). For the right pan-tilt unit, we have the unit centre \( T' \), pan angle \( a' \) and the title tilt \( \beta' \) as well as the angle \( \gamma' \).

The left and right pan-tilt units are moveable but only symmetrically left-right about the central axis S-W along the S-U axis in accordance with the requirement on the stereo baseline length change due to different photogrammetric precision requirement. Suppose the positions of \( T \) and \( T' \) are \( T(U_T, V_T, W_T) \) \( T'(U'_T, V'_T, W'_T) \), respectively, in the reference system \( S-UVW \). In general, we require the system to satisfy

\[
\begin{align*}
U_T &= -U_T = s, \\
V_T &= V'_T, \\
W_T &= W'_T \tag{12}
\end{align*}
\]

where \( s \) is a symbol denoting the distance from the left or right unit centre to the centre of the
stereo plate, which is about the half of the baseline length. Note that a primary difference of Photogrammetron I from comaison robots is that the baseline length is changeable and controlled by the system.

Although each of the left and right camera pan-tilt units has two angular freedoms, we distinguish between two general system modes: strong stereo mode versus weak stereo mode. On the strong stereo mode, two principal axes C-z and C'-z' must be maintained coplanar, and that plane is called the principal epipolar plane. The two principal axes C-z and C'-z' form two angles $\theta$ and $\theta'$ with the baseline CC', respectively. In the weak stereo mode, we do not require the two principal axes to be strictly coplanar, but the left and right camera should maintain overlapping views. We shall discuss the calibration for the two modes, respectively.

3 In-lab calibration of fixed parameters

In the geometric model described above, there are the following fixed relations.

1) The stereo plate is fixed on the top of the head, so the distance parameter $h$ is a constant;

2) The left and right units can only translate in one dimension, so the other two distance parameters $V_r, W_r$ and $V_r', W_r'$ are constants;

3) The left camera is fixed on the top of the left unit, so the translations and scaling $a, b, c, d$ as expressed in Eq. (6) are constant, which mediate the influence of the pan and tilt angles of the unit to the perspective centre.

The set of constant parameters is, therefore, defined as

$$(h, V_T, W_T, V'_T, W'_T, a, b, c, d)$$

The constant parameters $h, V_T, W_T, V'_T, W'_T$ can be measured through pure mechanical procedures. The constants $a, b, c, d$ are determinants of the perspective centre of the camera relative to the pan-tilt unit. They have to be determined by using control information such as control points in a laboratory set-up. However, the actual procedures for determining these constants can be the bundle adjustment by using the perspective equations which are available in photogrammetry literatures.

4 In-situ calibration for the strong stereo mode

In the strong stereo mode, for the simplicity of the geometry, we freeze the tilt freedom of the left and right camera units to absolute zero, so the two principal axes are coplanar with the $S$-$UV$ plane. The rest pan angle of the left or right unit is now denoted by $\theta$ and $\theta'$, respectively, as shown in Fig. 1, i.e.

$$\theta = \alpha, \theta' = \pi - \alpha', \beta = \beta' = 0$$

With the reference systems and geometric elements defined above, we can establish the stereo imaging equations. Take $O$-$XYZ$ as the global reference system. At any time $t$, an object point $P(x, y, z, t)$ is projected through the two cameras onto the left and right image points $p(x', y', f', t)$, $p'(x', y', f', t)$ on the left and right images $I, I'$, then the corresponding image values are $I(x, y, t), I'(x', y', t)$. The projective equation between $P$ and $p$ can be expressed as

$$P = \tilde{O}P = \tilde{O}S + \tilde{S}C + \lambda \tilde{C}P$$

where $\lambda$ is a scalar written in analytical form, then we have

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ h \end{bmatrix} + R_r R_s (+R_s R_r) \begin{bmatrix} U_c \\ V_c \\ W_c \end{bmatrix} + \lambda R_s R_r \begin{bmatrix} x - x' \\ y - y' \\ -f \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \\ h \end{bmatrix} + R_s R_r (+R_s R_r) \begin{bmatrix} x - x' \\ y - y' \\ -f \end{bmatrix}$$

Similarly we can derive the projective equation between $P$ and $p'$ for the right camera as

$$P = \tilde{O}P = \tilde{O}S + \tilde{S}C' + \lambda' \tilde{C}'P'$$

or in analytical form as

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ h \end{bmatrix} + R_r R_s (+R_s R_r) \begin{bmatrix} U_c \\ V_c \\ W_c \end{bmatrix} + \lambda' R_s R_r R' \begin{bmatrix} x' - x' \\ y' - y' \\ -f' \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \\ h \end{bmatrix} + R_r R_s (+R_s R_r) \begin{bmatrix} x' - x' \\ y' - y' \\ -f' \end{bmatrix}$$
where $\lambda'$ is a scalar and $R_f$ has the form of $R$, as defined in Eq. (7) with $a$ replaced by $n_0'$. Let

$$R = R_1R_2R_3R_4 = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

(19)

$$R' = R_1R_2R_3R_4 = \begin{bmatrix} r'_{11} & r'_{12} & r'_{13} \\ r'_{21} & r'_{22} & r'_{23} \\ r'_{31} & r'_{32} & r'_{33} \end{bmatrix}$$

(20)

$$\begin{cases} u \\ v \\ w \end{cases} = R \begin{cases} x - x_c \\ y - y_c \\ f \end{cases}$$

(21)

$$\begin{cases} u' \\ v' \end{cases} = R' \begin{cases} x' - x_c' \\ y' - y_c' \\ f' \end{cases}$$

(22)

then we have

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ h \end{bmatrix} + R_1R_2 \begin{bmatrix} U_c \\ V_c \\ W_c \end{bmatrix}$$

(23)

$$\begin{bmatrix} X_c' \\ Y_c' \\ Z_c' \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ h \end{bmatrix} + R_1R_2 \begin{bmatrix} U_c \\ V_c \\ W_c \end{bmatrix}$$

(24)

Eqs. (16) and (18) now can be rewritten as

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} + \lambda \begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

(25)

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_c' \\ Y_c' \\ Z_c' \end{bmatrix} + \lambda' \begin{bmatrix} u' \\ v' \end{bmatrix}$$

(26)

Eliminating the scalar $\lambda$ and $\lambda'$ from the above equations results in the collinearity equations

$$\begin{bmatrix} X - X_c \\ Y - Y_c \\ Z - Z_c \end{bmatrix} = \begin{bmatrix} r_{11} & r_{21} & r_{31} \\ r_{12} & r_{22} & r_{32} \\ r_{13} & r_{23} & r_{33} \end{bmatrix} \begin{bmatrix} x' \end{bmatrix}$$

(27)

$$\begin{bmatrix} X - X_c \\ Y - Y_c \\ Z - Z_c \end{bmatrix} = \begin{bmatrix} r'_{11} & r'_{21} & r'_{31} \\ r'_{12} & r'_{22} & r'_{32} \\ r'_{13} & r'_{23} & r'_{33} \end{bmatrix} \begin{bmatrix} y' \end{bmatrix}$$

(28)

For each object point we have four collinearity equations at time $t$. Note that there are only seven free parameters, which are controlled by the system:

$$x = (\omega, \phi, \theta, \theta', s, f)^T$$

(31)

Applying Eqs. (6), (7), (8), (9), (11), (19), (20), (23), (24) to equations (27)-(30), we obtain the functional forms of $x, y, x', y'$:

$$x = F(\omega, \phi, \theta, \theta', s, f; X, Y, Z)$$

(32)

$$y = G(\omega, \phi, \theta, \theta', s, f; X, Y, Z)$$

(33)

$$x' = F'(\omega, \phi, \theta, \theta', s, f'; X, Y, Z)$$

(34)

$$y' = G'(\omega, \phi, \theta, \theta', s, f'; X, Y, Z)$$

(35)

For target tracking, we must assume that the object points are also moving and every free parameter is also changing with time, so the collinearity equations should be written as

$$x(t) = F(\omega, \phi, \theta, \theta', s; f_1 X(t), Y(t), Z(t))$$

(36)

$$y(t) = G(\omega, \phi, \theta, \theta', s; f_1 X(t), Y(t), Z(t))$$

(37)

$$x'(t) = F'(\omega, \phi, \theta, \theta', s; f_1' X(t), Y(t), Z(t))$$

(38)

$$y'(t) = G'(\omega, \phi, \theta, \theta', s; f_1' X(t), Y(t), Z(t))$$

(39)

However, for system calibration, we assume that a number of control points exist in the surveillance area, and that they are either man-made or extracted feature points but all fixed still. For each control point, we have four collinearity equations, which are continu-
ous in time $t$

\begin{align}
 x(t) &= F(\omega, \varphi, \theta, \theta', s, f; X, Y, Z; t) \\
 y(t) &= G(\omega, \varphi, \theta, \theta', s, f; X, Y, Z; t) \\
 x'(t) &= F'(\omega, \varphi, \theta, \theta', s, f'; X, Y, Z; t) \\
 y'(t) &= G'(\omega, \varphi, \theta, \theta', s, f'; X, Y, Z; t)
\end{align}

There are basically two approaches for solving these equations to determine the seven free parameters which may change continuously in time.

The first approach uses $n > 2$ control points to form $4n$ collinearity equations of the forms (40)-(43), then the seven free parameters at any time point $t$ are obtained. The actual procedure is similar to the bundle adjustment in analytical photogrammetry\textsuperscript{23}, but with the particular parameter set of Eq. (31). We shall not delve into the details of this approach as the bundle adjustment is well known in photogrammetry, and this particular bundle adjustment can be developed in a similar way.

The second approach is on the basis of the first approach, and also takes account of the continuity of the parameter variables and system dynamics and take the system reading of these parameters as observations to the parameters themselves. The state transition equations and the observation equations of the Kalman filtering\textsuperscript{32} are written as

\begin{align}
 x(t_k) &= \Phi(t_k, t_{k-1}) x(t_{k-1}) + \Gamma(t_{k-1}) w(t_{k-1}) \\
 z(t_k) &= \Psi(t_k) x(t_k) + v(t_k) \\
 \text{or by simplified notations they become} \\
 x_k &= \Phi_k x_{k-1} + \Gamma_k w_{k-1} \\
 z_k &= \Psi_k x_k + v_k \\
 \text{where} x(t) \text{is the seven-dimensional parameter vector as defined by Eq. (31) at time} t, \text{also called the state vector of the system;} k \text{is the integer index of time and satisfies} \\
 -\infty < k - 1 < t_{k-1} < t_k < t_{k+1} < k < \infty
\end{align}

$w(t)$ is the $m$-dimensional dynamic noise vector, $z(t)$ is the $l$-dimensional observation vector, $l \leq 7 + 4n$, which includes the system readings of the free parameters and image coordinates $(x, \varphi, \theta, \theta', s, f)$ of visible control points. Note that not every free parameter or every control point is visible, the observation vector may be incomplete data. And $\Phi(t, \tau)$ is a $7 \times 7$ non-singular matrix called the state transition matrix of the system, $\Gamma(t)$ is a $7 \times m$ matrix called the dynamic noise matrix, $\Psi(t)$ is a $7 \times 7$ matrix called the observation matrix. Moreover, $\Phi(t, \tau)$ has the following properties:

1) $\Phi(t, \tau) = I$ (where $I$ is an identity matrix)
2) $\Phi^{-1}(t_k, t_{k-1}) = \Phi(t_{k-1}, t_k)$
3) $\Phi(t_k, t_{k-2}) = \Phi(t_k, t_{k-1}) \Phi(t_{k-1}, t_{k-2})$

The observation Eq. (45) include the linearized version of the collinearity Eqs. (40)-(43) as well as the additional observation equations of the system readings for the parameters $x(t)$. We shall not delve into the detailed form of the state transition Eq. (44) and the observation Eq. (45).

Let $\hat{x}_k$ denote the estimate of $x(t)$ at time $t_k$, and $\tilde{x}_k$ denote the error of estimation. Supposing the estimate $\hat{x}_k$ is a linear function of the observation $z$, the linear least square estimation is achieved under the following criterion

\begin{align}
 \min_{\hat{x}_k} E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T] = E[z_k x_k^T]
\end{align}

Supposing $k$ observations $z_1, z_2, \ldots, z_k$ to the seven-dimensional linear dynamic system of Eq. (44) have been made through the $l$-dimensional linear observation system of Eq. (45) from time 1 to time $k$. According to these $k$ observation data, we can estimate the system state $\hat{x}_k$ at time $k$, and the actual estimation procedure has a particular form of Kalman filtering as

\begin{align}
 \hat{x}_k &= \hat{x}_{k-1} + K_k (z_k - \Psi_k \Phi_k \hat{x}_{k-1}) \\
 \text{where} K_k \text{is called the weight matrix or gain matrix and defined by the coefficient matrices of the state transition Eq. (44) and the observation Eq. (45) as well as the stochastic properties of the noises} w_k, v_k, \text{. We shall not delve into the detailed form of} K_k \text{and further details of the estimation procedure due to the space limitation.}
5 In-situ calibration for the weak stereo mode

In the weak stereo mode, each of the left or right pan-tilt unit has two angular freedoms $\alpha$, $\beta$ (or $\alpha'$, $\beta'$ for the right camera) and the principal axes of the left camera and the right one are not required to be coplanar. In this mode, the free parameter vector consists of nine free parameters which may change in time,

$$\mathbf{x} = (\omega \phi a \beta a' \beta' s f f)^T$$ (54)

There are two approaches for calibration in such a weak stereo mode; the first is a joint solution for estimating all the nine parameters simultaneously through a particular form of Kalman filtering as described in the previous section; the second is to estimate the absolute orientation and interior orientation for each camera independently through control points. Still in the second approach, the continuity and dynamics of the system state parameters can be exploited by a Kalman filtering mechanism.

6 Experimentation

In our experiment, we adopted a colour CCDX ZOOM HIACHI video camera without any preengagement parameter. The calibration steps are described as follows.

1) According to the known control points and the corresponding image points from the image segmentation, calculate the interior/exterior orientation element and affixation parameter of camera, with the use of DLT.

2) Use these values as the initial values to attach parameter self-calibration bundle adjustment.

3) On the basis of Kalman filtering described above, calibrate the free parameter continuously and dynamically.

The concrete work is in process.

7 Conclusions

In this paper, a theory of geometric calibration of intelligent photogrammetron is proposed upon a geometric model of photogrammetron. Two system-operating modes are distinguished: the strong stereo mode versus the weak stereo mode. In the strong stereo mode, the free parameter vector comprises seven parameters, while in the weak stereo mode each of the left or right pan-tilt unit has its own pan and tilt angular freedoms. A pure photogrammetry solution is a particular bundle adjustment using fixed control points. However, the general solution is a particular Kalman filtering which builds on the bundle adjustment but extends to exploiting the continuity and dynamics of system motion. The theory proposed here is quite general, but any actual implementation has to take account of the actual physical structure and control mechanisms of the photogrammetron system.

References

1 Pan H P (2002) Concepts and initial design of intelligent stereo photogrammetron. Journal of Surveying and Mapping, 31(4):310-315
2 Wang Z Z (1990) Principles of photogrammetry with remote sensing. Beijing: Publishing House of Surveying and Mapping,
3 Kalman R E (1960) A new approach to linear filtering and prediction problems. Transactions of the ASME, Journal of Basic Engineering, 35-46
4 Pan H P (2002) A basic theory of intelligent photogrammontron. Photogrammetric Computer Vision ISPRS Commission III Symposium, Graz.