DYNAMIC MODELING AND GEOREGISTRATION OF AIRBORNE VIDEO SEQUENCES

H. J. THEISS[*], C. LEE[*], E. M. MIKHAIL[*], M. R. Vriesenga[**]
[*]Purdue University, USA
Geomatics Engineering Department
theiss, changno, mikhail@ecn.purdue.edu
[**]BAE Systems, USA
mvriesen@san.rr.com

IC Working Group V/III

KEY WORDS: Calibration, Dynamic processes, Image registration, Image sequence, Photogrammetry

ABSTRACT

Rigorous sensor and dynamic modeling techniques are required if spatial information is to be accurately extracted from video imagery. First, a math model for an uncalibrated video camera and a description of a bundle adjustment with added parameters, for purposes of general block triangulation, is presented. The next topic involves the application of invariance-based techniques, with constraints, to derive initial approximations for the camera parameters. Finally, dynamic modeling using the Kalman Filter is discussed.

1 INTRODUCTION

Unmanned aircraft vehicles (UAV's) with video cameras on board are becoming a popular reconnaissance tool, particularly for military applications. Video cameras are generally uncalibrated and inherently contain significant lens distortions that require careful mathematical modeling. Unlike metric cameras, there are no fiducial marks visible on video imagery to allow the precise location of image coordinates with respect to the principal point. First, the application of a bundle adjustment with added parameters to perform simultaneous photogrammetric georegistration of a general block of video frames is presented. Next, invariance-based techniques to compute initial estimates for camera parameters are given. Finally, dynamic modeling of video sequences using the Kalman Filter is discussed.

2 SIMULTANEOUS PHOTOGRAMMETRIC GEOREGISTRATION

2.1 Rigorous Sensor Model

Rigorous sensor modeling of a video sequence consists of three major parts: 1) object-to-image transformation from the ground-space coordinate system (X,Y,Z) to the image-space coordinate system (x,y,z) modeled by collinearity as a function of six exterior orientation (EO) parameters; 2) the transformation from raw observed line and sample image coordinates in a 2D pixel array to the image-space coordinate system (x,y,z) using interior orientation (IO) parameters to model several types of systematic errors; and 3) platform modeling which considers the stochastic relationship among camera parameters of adjacent frames, the focus of Section 4.

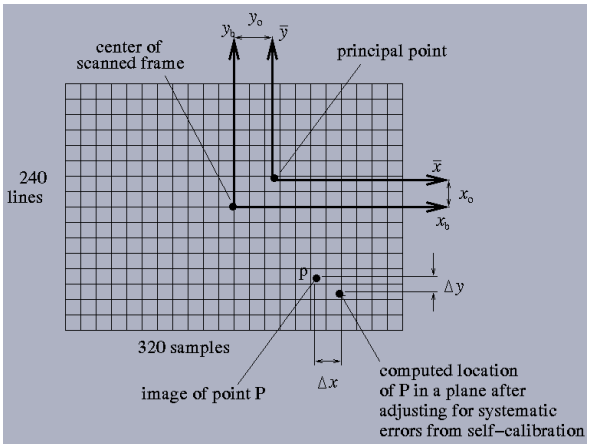

Figure 1. Video interior orientation

The collinearity condition equations are a function of the following parameters (see Figure 1):

- x_o,y_o,c are the elements of interior orientation, the principal point offsets (x_o,y_o) and the camera principal distance,
- x_b,y_b observed image coordinates,
- X,Y,Z ground coordinates corresponding to the image points,

M	camera orientation matrix (function of these angles, ω, φ, κ),
X_L,Y_L,Z_L	camera location,
Δc	principal distance correction parameter,
K_1,K_2,K_3	three radial lens distortion parameters,
p_1, p_2	two decentering lens distortion parameters, and
b_1, b_2	two in-plane distortion parameters (skew and scale difference).

2.2 Bundle-Block Adjustment Algorithm

For each image point that appears in a video frame, we write two collinearity condition equations of the form $F(l, \underline{x})=0$, where l is the vector of observables (image point coordinates, x_b, y_b, for each image point) and $\underline{x}$ is the vector of all unknown parameters (IO elements and EO elements for each camera, and object point ground coordinates, X, Y, Z, for each ground point). The linearized set of collinearity equations takes the general form $Av + B\Delta = f$ (Mikhail 1976).

A unified least squares technique is implemented in order to accommodate the recovery of different sets of unknown parameters, thus allowing for the solution of various photogrammetric problems. The unified technique allows for the incorporation of a priori information about the parameters, $\underline{x}$; i.e, it treats the parameters as observations with a priori covariance matrix, Σ_{xx}. Implementation of the unified least squares adjustment technique with the general least squares model can also be found in (Mikhail 1976).

The parameter, c, is fixed and taken as a constant in the adjustment. The variables x_b and y_b are considered observables in the adjustment. The ground point parameters, X,Y,Z, are considered free adjustable parameters with very large covariances for the unknown pass points. The X,Y,Z for control points however have associated covariances that represent the degree of accuracy to which they are known. All of the six exterior orientation (EO) parameters, X_L, Y_L, Z_L, ω, φ, κ, are free adjustable parameters with associated high a priori covariances. Only a subset of the interior orientation (IO) parameters, x_o, y_o, Δc, K_1, K_2, K_3, p_1, p_2, b_1, b_2, can actually be recovered in the adjustment since the normal equations become unstable due to high correlations among them and between them and the EO parameters. The IO parameters to be recovered are free adjustable parameters with associated apriori covariances.

2.3 Experiments

The (VA Hospital) data set was used: H = 900 m, GSD (nadir) = 3 m, t = 30°, f = 7.4 mm, format 320×240 pixels, pixel = 20μm, $\sigma_{control}$ = 0.15 m, σ_{image} = 0.5 pixel. The three cases were: 1) single frame resection, 2) three frames from the same strip overlapping approximately 60%, 3) a block of six frames, three from each of two convergent strips.

Experiment 1:
Photogrammetric resection for a single video frame, Frame 41. Several different sets of recovered interior orientation (IO) parameters are tested for comparison. The first ten cases are with 33 control points and 8 check points, while the last two cases are with only 6 control points and 35 check points. The control and check point RMS are tabulated in Table 1. In check point computation, note that the Z coordinate is fixed to its known value, while the X and Y coordinates are computed using the inverse form of the collinearity equations.

Case: IO Parameters	No. Control Points	No. Check Points	Control Point RMS (m)			Check Point RMS (m)		
			X	Y	plan.	X	Y	plan.
0: x_o, y_o,Δc	33	8	2.01	2.31	3.06	2.03	1.93	2.80
1: x_o, y_o,K_1	33	8	1.38	1.37	1.95	1.41	1.82	2.30
2: x_o, y_o,Δc,K_1	33	8	1.28	1.39	1.89	1.31	1.83	2.25
3: x_o, y_o,Δc,K_1, p_1, p_2	33	8	1.30	1.33	1.86	1.34	1.89	2.31
4: x_o, y_o,Δc,K_1, b_1, b_2	33	8	1.30	1.36	1.88	1.37	1.87	2.31
5: x_o, y_o,Δc,K_1, p_1, p_2, b_1, b_2	33	8	1.30	1.33	1.87	1.33	1.85	2.28
6: x_o, y_o,Δc,K_1,K_2	33	8	1.28	1.39	1.89	1.30	1.83	2.24
7: x_o, y_o,Δc,K_1,K_2,K_3	33	8	1.27	1.39	1.88	1.25	1.77	2.17
8: x_o, y_o,Δc,K_1,K_2, p_1, p_2, b_1, b_2	33	8	1.27	1.37	1.87	1.28	2.07	2.44
9: x_o, y_o,Δc,K_1,K_2,K_3, p_1,p_2, b_1,b_2	33	8	1.32	1.30	1.85	1.51	2.40	2.83
1: x_o, y_o,K_1	6	35	0.36	0.21	0.42	1.97	2.19	2.95
2: x_o, y_o,Δc,K_1	6	35	0.19	0.18	0.26	1.88	2.13	2.84

Table 1. Resection Results for Single Uncalibrated Video Frame

Comments: 1) adding IO parameters improves the control point RMS, but not for check points; 2) recovering a radial lens distortion parameter, K_1, is necessary for resection with an uncalibrated video camera (compare case 1 to case 0); 3) recovering Δc improves the resection results (compare case 2 to case 1); 4) estimating p_1, p_2, b_1, b_2 does not improve check point RMS (compare cases 3-5 to case 2); and 5) recovering K_2 and K_3 improved the results a little. For a practical number of control points, it is not feasible to use more parameters than those in case 2.

Experiment 2:
The second experiment involved the triangulation of three video frames with approximately 60% forward overlap. Since the geometry of the intersection of rays was poor for this case, the value of the *Z* coordinate was fixed while the *X* and *Y* coordinates of check points were computed. Six cases of recovered IO parameters were tested and the RMS results are tabulated in Table 2. For this experiment, the best and most practical choice of sensor model parameters appears to be case 2, which recovers x_o, y_o, Δc, and K_1.

Case: IO Parameters	Control Point RMS (m)			Check Point RMS (m)		
	X	Y	plan.	X	Y	plan.
1: x_o, y_o, K_1	1.39	1.06	1.74	2.82	2.70	3.91
2: $x_o, y_o, \Delta c, K_1$	0.91	0.78	1.20	1.72	2.43	2.97
3: $x_o, y_o, \Delta c, K_1, p_1, p_2$	0.88	0.52	1.02	2.83	2.71	3.91
4: $x_o, y_o, \Delta c, K_1, b_1, b_2$	0.90	0.48	1.01	2.54	2.67	3.69
6: $x_o, y_o, \Delta c, K_1, K_2$	0.90	0.79	1.20	1.72	2.43	2.97
7: $x_o, y_o, \Delta c, K_1, K_2, K_3$	0.90	0.80	1.20	1.71	2.43	2.97

Table 2. Triangulation Results: 3 Frames (1 strip), 31 pass points, 7 control points, 15 check points

Experiment 3:
A block of six video frames, with 60% overlap and 100% sidelap, is used in a simultaneous triangulation. This experiment compares the use of a relatively large number of pass points, 19, to the use of only 5 pass points. Note that results from these two point configurations are shown in different columns of the same Table 3. Therefore, only check point results are shown. For this relatively large block, check point RMS results are consistent for different parameterization cases. Note that the check point RMS results become significantly worse, especially in the Z direction, when using relatively much fewer pass points.

Case: IO Parameters	Check Point RMS (m)							
	19 Pass Points				5 Pass Points			
	X	Y	Z	radial	X	Y	Z	Radial
1: x_o, y_o, K_1	1.21	2.18	2.44	3.49	1.17	2.64	3.82	4.79
2: $x_o, y_o, \Delta c, K_1$	1.05	2.17	2.58	3.53	1.01	2.52	3.53	4.45
3: $x_o, y_o, \Delta c, K_1, p_1, p_2$	1.04	2.16	2.56	3.50	1.01	2.46	3.43	4.34
4: $x_o, y_o, \Delta c, K_1, b_1, b_2$	1.04	2.16	2.58	3.52	1.04	2.50	3.86	4.71
5: $x_o, y_o, \Delta c, K_1, p_1, p_2, b_1, b_2$	1.04	2.16	2.48	3.45	1.06	2.47	3.96	4.79
6: $x_o, y_o, \Delta c, K_1, K_2$	1.05	2.17	2.58	3.53	1.03	2.53	3.51	4.45
7: $x_o, y_o, \Delta c, K_1, K_2, K_3$	1.06	2.22	2.51	3.51	Did not converge.			
8: $x_o, y_o, \Delta c, K_1, K_2, p_1, p_2, b_1, b_2$	1.06	2.16	2.59	3.53	Did not converge.			

Table 3. Triangulation Results: 6 Frame Video Block (2 convergent strips); 7 control points, 14 check points

3 INVARIANCE-ASSISTED VIDEO TRIANGULATION

Since the techniques in Section 2 are non-linear and require good parameter initial approximations, the invariance techniques to be described in this section are practical, and can be applied as a first step in any photogrammetry-based georegistration algorithm. The purpose of invariance is to develop functional relationships such that the equations are linear with respect to the parameters. Unlike photogrammetry, however, the parameters used in an invariance formulation do not usually have a one to one correspondence with the physical characteristics of the camera being modeled or of the camera's location and orientation.

3.1 Linear Formulation – The *P* Matrix

The first step in recovering the camera parameters from a single video frame is to estimate 11 of the 12 elements of the 3×4 camera transformation matrix, *P*, which projectively relates ground coordinates to image coordinates as $[x \quad y \quad 1]^T \approx P[X \quad Y \quad Z \quad 1]^T$ in which (*x*,*y*) are image point coordinates, and (*X*,*Y*,*Z*) are ground point coordinates. The "≈" implies equality up to a scale factor. To cancel out the scale factor, we can divide the first and second equations by the third equation, respectively, noting that $p_{34} = 1$.

$$F_{xj} = x_j - \frac{P_1[X_j \quad Y_j \quad Z_j \quad 1]^T}{P_3[X_j \quad Y_j \quad Z_j \quad 1]^T} = 0, \qquad F_{yj} = y_j - \frac{P_2[X_j \quad Y_j \quad Z_j \quad 1]^T}{P_3[X_j \quad Y_j \quad Z_j \quad 1]^T} = 0 \tag{1}$$

where: *m* is the number of points used in the adjustment, $m \geq 6$,
j is the point number, *j* = 1, 2, . . ., *m*,
P_i is the i[th] 1×4 row vector of the *P* matrix, *i* = 1, 2, 3, and
(x_i, y_i) are the observed image coordinates.

A linear pseudo least squares solution can be applied to Equations (1) to solve for the p_{ij}, by minimizing errors in the linear equations obtained by clearing fractions. If a rigorous refinement is desired, then least squares is applied directly to Equations (1). Since Equations (1) are nonlinear with respect to the unknowns, linearization in the form $v + B\Delta = f$ (Mikhail, 1976) is required using the estimates obtained from the linear least squares solution as initial approximations.

3.2 Estimation of Photogrammetric Camera Parameters from *P*

It can be shown that the camera transformation matrix, *P*, can be partitioned as follows (Barakat and Mikhail, 1998):

$$P = k\left[\underset{3\times3}{AM} \quad | \quad -\underset{3\times1}{AMT}\right], \qquad T = \begin{bmatrix} X_L \\ Y_L \\ Z_L \end{bmatrix}, \qquad A = \begin{bmatrix} 1 & -b_2 & \frac{-x_o}{c}(1+b_1) \\ 0 & 1+b_1 & \frac{-y_o}{c}(1+b_1) \\ 0 & 0 & \frac{-1}{c}(1+b_1) \end{bmatrix} \tag{2}$$

where *M* is the orthogonal camera orientation matrix (function of ω, φ, κ) **,**
x_o, y_o are the principal point offset parameters,
c is the principal distance, and
b_1, b_2 are the two in-plane distortion parameters; i.e., scale difference and skew, respectively.

The matrices *M* and *A* are obtained applying the QR decomposition, which decomposes a square matrix into an upper triangular matrix and an orthogonal matrix. The five interior orientation elements can be extracted from *A*, the camera location can be extracted from *T*, and the orientation angles can be extracted from *M* as shown in (Barakat and Mikhail, 1998).

3.3 Constraints Among Elements of *P*

In cases when it is known that, for all practical purposes, the pixels are square and the *x* and *y* axes intersect at a right angle, we can write two constraint equations to reduce the number of independent unknowns from 11 to 9. Enforcing the fact that the *M* matrix must be orthogonal, and b_1=b_2=0 in Equation (2) leads us to arrive at two such constraint equations, G_1 and G_2. If, in addition to the former two constraints, we have good estimates of the principal point offsets or principal distance such as from camera calibration, then up to three additional constraints may be written, G_3, G_4 and G_5; i.e., one for each of the three known constants x_o, y_o, and *c*. These three constraints would reduce the number of independent unknowns further, to 6. The five constraint equations, whose detailed derivation can be found in a technical report, (Theiss and Mikhail, 1999), are:

$$G_1 = (R_1 R_3^T)(R_2 R_3^T) - (R_1 R_2^T)(R_3 R_3^T) = 0 \quad (3a)$$

$$G_2 = (R_1 R_1^T)(R_3 R_3^T) - (R_1 R_3^T)^2 - (R_2 R_2^T)(R_3 R_3^T) + (R_2 R_3^T)^2 = 0 \quad (3b)$$

$$G_3 = (R_1 R_3^T) - (R_3 R_3^T)(x_o{}^o) = 0 \quad (3c)$$

$$G_4 = (R_2 R_3^T) - (R_3 R_3^T)(y_o{}^o) = 0 \quad (3d)$$

$$G_5 = (R_1 R_1^T)(R_3 R_3^T) - (R_1 R_3^T)^2 - (c^o)^2 (R_3 R_3^T)^2 = 0 \quad (3e)$$

where $x_o{}^o$, $y_o{}^o$, and c^o are the apriori values of the principal point offsets and principal distance, respectively. Note that any combination of the constraints (Equations 3a-e) may be enforced depending on the photogrammetric implications of the problem.

3.4 Fundamental Matrix (*F*) Relationship for a Pair of Video Frames

The fundamental matrix directly relates the image coordinates of 3D object points that appear on two images. The 3 by 3 *F* matrix has eight unknown parameters since it is determinable up to a scale factor, and its (3,3) element is set equal to unity. In fact there are seven independent parameters, since *F* is of rank two and its determinant must be zero. Once solved for, the *F* matrix can be factorized into two relative camera matrices, P_{r1} and P_{r2} (Barakat and Mikhail, 1998). Then the projective model coordinates, $(X,Y,Z,1)_r$, can be computed as a function of the image coordinates of a point on two images and their associated relative camera transformation matrices.

For uncalibrated cameras, the model coordinates computed using relative camera transformation matrices are in a 3D non-conformal system. Given the 3D model coordinates, the 15 elements of a non-singular 4×4 projective transformation matrix, *H*, are computed. Since three equations per point can be written, a minimum of 5 points is required to solve for the 15 elements of *H*. (Note that the (4,4) element of *H* is set to unity.) With more than 5 points, a linear least squares solution is applied. The 3D projective transformation *H* is from projective ground space to projective model space. Once solved for, the *H* matrix may be used to compute either absolute ground coordinates, or the absolute camera transformation matrices. Finally, the photogrammetric camera parameters can be extracted from the camera transformation matrix, *P*, using the techniques discussed in Section 3.2.

3.5 Experiments

3.5.1 Single Video Frame Camera Parameter Recovery. Experiments were run on two different video frames to test the ability to estimate camera parameters, as a function of image and ground coordinates only, to be used as initial approximations for rigorous photogrammetry. Both frames came from the VA Hospital data set. The first frame, 5800_17, was taken at 760 meters above mean terrain elevation with estimated principal distance and GSD (at nadir) of 16.7 mm and 1.3 meters, respectively. The second frame, 6100-68, was taken at 950 meters above terrain with estimated principal distance and GSD of 19.2 mm and 1.4 meters, respectively. Both frames had nominal side-look angles of 45 degrees from nadir.

The steps involved are: 1) compute the 11 elements of the P matrix using linear least squares, as described in Section 3.1; 2) apply the non-linear least squares with the first two constraints on the 11 elements of P, using the results of Step 1 as initial approximations, as described in Section 3.3; 3) estimate the real photogrammetric camera parameters from the P matrix, i.e., the physical 6 exterior orientation and 3 interior orientation elements, as described in Section 3.2; 4) use these 9 camera parameters as initial approximations in a rigorous photogrammetric resection; 5) use the principal distance estimated from Step 4 and the fair assumption that $x_o = y_o = 0$ and 3 additional constraints (a total of 5 constraints) to estimate only the 6 exterior orientation parameters; and 6) Compute the RMS values using some check points for each method to assess the performance.

The check point results for each of the video frames are tabulated in Table 4. The X, Y, and Z ground coordinates of control points were manually extracted from a triangulated stereopair of frame images on the digital photogrammetric workstation.

3.5.2 Two-Frame Video Camera Parameter Recovery. The pair has nearly 100% overlap, taken from different flight lines. The *F*-matrix technique was applied as described in Section 3.4. The resulting camera transformation

matrices, *P*, for each video frame which were used together to compute the X, Y, and Z ground coordinates of known check points; the RMS results are shown in the first line of Table 5.

Method	No. of Unkn.	Check Point RMS (m)					
		Frame 5800-17, 8 control, 7 check			Frame 6100-68, 8 control, 8 check		
		X	Y	Planim.	X	Y	Planim.
Inv	11	0.70	1.10	1.28	2.90	1.58	3.29
Inv	9	0.67	0.94	1.16	3.51	1.95	3.99
Photog.	9	0.67	0.94	1.16	3.51	1.95	4.02
Inv	6	1.16	0.70	1.34	3.60	2.90	4.63

Table 4. VA Hospital: Single Video Frame RMS

Case	Check Point RMS (m)		
	X	*Y*	*Z*
F-Matrix	1.79	4.78	1.72
Rigorous Photogr.	0.84	2.94	2.62

Table 5. VA Hospital, Frames 7430_75 and 7030_70; Check Point RMS

The camera parameters recovered from the *P*'s were then used as input in a rigorous photogrammetric adjustment, and check point ground coordinates were computed for the same points. Ten camera parameters were recovered for each video frame, including 4 interior orientation (IO) parameters in addition to the 6 exterior orientation parameters. The four IO parameters included the principal point offsets, x_o and y_o, the principal distance, c, and a radial lens distortion parameter, K_1. Although invariance provides good camera parameter initial approximations as input to the rigorous solution, rigorous photogrammetry with an additional parameter for lens distortion provides better results.

4 DYNAMIC MODELING

4.1 Kalman Filter Estimation

Information obtained from the processing of previous video frames can be used to essentially constrain the solution of a current frame to have reasonable parameter estimates that are consistent with its neighboring frames. In other words, there is a stochastic relationship and high correlation between at least some of the sensor model parameters of the current and neighboring video frames. The Gauss-Markov (GM) process is an example of such a stochastic model. A random process is a collection of functions of time. More specifically, a first order Markov process is a continuous random process that satisfies the criterion that the current state is dependent only on the state at the previous point in time.

Kalman Filtering is a useful technique that allows implementation of sequential least squares adjustment while simultaneously allowing the enforcement of a stochastic process. It does this in the form of a state transition matrix, and will be described in more detail in this section.

The Kalman Filter equations can be written as follows for the i^{th} frame in a sequence of video images (Brown and Hwang 1997),

$$\Delta_i = \Delta_i^- + K_i(f_i - B_i\Delta_i^-), \quad K_i = Q_{\mathrm{xx}_i}^- B_i^T (B_i Q_{\mathrm{xx}_i}^- B_i^T + Q_{e_i})^{-1} \tag{4a,b}$$

where K_i is the Kalman Gain, and Δ_i^- and $Q_{\mathrm{xx}_i}^-$ are the a priori Δ and Q_{xx}, respectively, computed at the end of processing the previous frame.

The following equations are for the covariance matrix of the updated state vector estimate, and estimates (projected ahead to the next frame) for the state vector elements and associated covariance matrix:

$$Q_{\mathrm{xx}_i} = (I - K_i B_i) Q_{\mathrm{xx}_i}^-, \quad \Delta_{i+1}^- = \phi_i \Delta_i, \quad Q_{\mathrm{xx}_{i+1}}^- = \phi_i Q_{\mathrm{xx}_i} \phi_i^T + Q_{ww} \tag{5a-c}$$

where ϕ_i is the $(u \times u)$ state transition matrix that transforms the current camera parameter state vector of frame *i* to its predicted state vector of frame (*i*+1). If this sequential approach is to be modeled as a first order GM process, then ϕ_i is a diagonal matrix that contains e^{-s_q} on each diagonal element, where s_q is a correlation factor unique to each parameter *q* (note $0 \le s_q \le 1$). The value of s_q is close to 0 for parameters that are highly correlated with their values at the previous frame, while for the contrary s_q is close to unity (Lee, 1999).

These Kalman Filtering equations provide the optimal parameter estimates for frame *i*, given all information from the first frame through the current frame, written $\Delta_{i|i}$. Therefore, the only frame whose parameters are estimated based on observations from all frames is the last frame, *n*. In order to obtain the optimal parameter estimate at an intermediate frame *i* based on measurements from all *n* frames, $\Delta_{i|n}$, we need to apply a backward smoothing process following the Kalman Filtering.

If smoothing is to be applied later, then for each frame in the Kalman Filtering algorithm it is necessary to save the a priori and a posteriori parameter estimates and their associated covariance matrices. At the completion of the forward sweep, the final computations of Equations (4a) and (5a) result in $\Delta_{n|n}$ and $Q_{\mathrm{xx}_{n|n}}$, respectively. Proceeding with the backward sweep, updated estimates, the smoothing gain, and the associated covariance matrix can be computed as follows,

$$\Delta_{i|n} = \Delta_{i|i} + K_{s_i}\left(\Delta_{i+1|n} - \Delta_{i+1|i}\right), \quad K_{s_i} = Q_{xx_{i|i}}\phi^T_{i+1|i}Q^{-1}_{xx_{i+1|i}}, \quad Q_{\mathrm{xx}_{i|n}} = Q_{\mathrm{xx}_{i|i}} + K_{s_i}\left(Q_{\mathrm{xx}_{i+1|n}} - Q_{\mathrm{xx}_{i+1|i}}\right)K_{s_i}{}^T \qquad \text{(6a-c)}$$

4.2 Experiments

Video sequence 7430 from the VA Hospital was used to test the developed Kalman Filtering algorithm. The flying height for strip 7430 was approximately 853 meters above ground level and the side-look angle was 30 degrees from nadir. The principal distance and ground sample distance (GSD) at nadir were estimated to be 354 pixels (approximately 7 mm) and 2.8 meters, respectively. The unknown camera parameters for each frame include the 6 EO parameters and 4 IO parameters. The IO parameters are the principal distance, principal point offsets, and one radial lens distortion coefficient; i.e., c, x_o, y_o, and K_1.

Since there was no GPS data available, known ground points were used to control the triangulation of the video sequence. The coordinates of the ground control points and check points were extracted from a controlled reference image base (CRIB), which consists of an orthophoto and its co-registered digital elevation model (DEM). This experiment uses 60 consecutive frames, 1 through 60, of strip 7430. There are 15 manually measured pass points with as many as possible measured on each frame. Control consists of 17 known ground points on frames 1 and 60 only; Figure 2 shows Frame 1. Since the geometric intersection between any two intermediate frames of the 60 frame strip would be poor for computing check points, the ground coordinates of the pass points themselves were evaluated simultaneously with the triangulation. Therefore, the check point differences for this experiment consist of the differences between the computed and the known coordinates of the15 pass points.

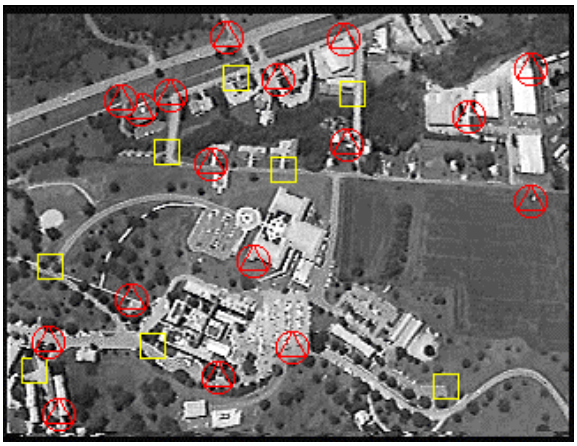

Figure 2. VA Hospital, Strip 7430, Frame 1

For all cases in this experiment, Q_{xx} was filled using the following a priori standard deviations: 100 meters for X_L, Y_L, Z_L and the X, Y, Z ground coordinates of all pass points; 0.1 meters for all ground coordinates of control points; 90 degrees for ω, φ, κ; 10 pixels for the principal point offset and principal distance; and 0.01 pixel^{-3} for K_1. The standard deviations for the white noise, or square root of the diagonal elements of Q_{ww} corresponding to the EO camera

parameters only, started with 100 meters for X_L, Y_L, Z_L and 90 degrees for ω, φ, κ and were decreased by a factor of 10 for each of the eight cases from (2a) through (2h). Thus, the standard deviations for X_L, Y_L, Z_L and for ω, φ, κ were: 100m and 90deg for case (2a), 10m and 9deg for (2b), 1m and 0.9deg for (2c), 0.1m and 0.09deg for (2d), 1cm and 30sec for (2e), 1mm and 3sec for (2f), 0.1mm and 0.3sec for (2g), and 0.01mm and 0.03sec for (2h). The check point results for cases (2a) through (2h) are tabulated in Table 6.

Case	RMS (m)				Bias (m)				Standard Deviation (m)			
	X	Y	Z	radial	X	Y	Z	radial	X	Y	Z	radial
2a	3.55	2.41	9.10	10.06	-2.01	-1.15	-6.20	6.62	2.93	2.12	6.65	7.57
2b	3.55	2.42	9.16	10.11	-2.04	-1.17	-6.28	6.71	2.91	2.12	6.66	7.57
2c	3.32	2.37	8.30	9.25	-1.98	-1.27	-5.99	6.44	2.67	1.99	5.75	6.64
2d	2.92	2.44	7.13	8.08	-1.76	-1.53	-5.38	5.87	2.33	1.91	4.68	5.56
2e	2.94	2.62	6.98	8.02	-1.54	-1.62	-5.06	5.53	2.51	2.06	4.80	5.80
2f	3.21	3.17	8.32	9.47	-1.57	-1.34	-6.95	7.25	2.80	2.88	4.57	6.08
2g	3.42	3.34	8.71	9.94	-1.62	-1.30	-7.33	7.62	3.02	3.08	4.71	6.39
2h	3.43	3.34	8.72	9.94	-1.62	-1.30	-7.33	7.62	3.02	3.08	4.71	6.39

Table 6. VA Hospital Check Point Results

To evaluate the triangulation accuracy, the root mean square (RMS), bias (signed mean), and standard deviation in *X*, *Y*, and *Z* of the check point differences was computed. The relationship between these three measures for each of *X*, *Y*, and *Z* can be written as follows:

$s = \sqrt{e^2 - b^2}$, where s = standard deviation, e = RMS error, and b = bias.

As the standard deviations of the white noise are decreased from case (2a) through (2d) the RMS of the check points generally decreases. However, as they are decreased further from case (2e) through (2h), the RMS's increase, especially in the *Y*, or flight, direction. The results suggest that the optimum choice of standard deviations of white noise to model the aircraft's trajectory for this sequence of frames is 0.1 meter for X_L, Y_L, Z_L and 0.09 degrees for ω, φ, κ.

5 CONCLUSIONS

Bundle adjustment with added parameters is effective for modeling video imagery, but care must be taken to recover the appropriate number and type of parameters depending on the geometry of the scene, and point configuration. Invariance-based solutions are helpful in obtaining initial approximations for camera parameters; however rigorous photogrammetry is superior with respect to accuracy. Dynamic modeling provides a means for real-time processing of video sequences, and incorporates the stochastic relationship among parameters of adjacent frames.

ACKNOWLEDGMENTS

This research is supported by the U.S. Air Force through a contract from BAE Systems (formerly Marconi, formerly GDE Systems). The authors thank Mr. Scott Stoneall and Dr. Will Snyder, from BAE Systems, for their contributions.

REFERENCES

Barakat, H. F., and Mikhail, E. M., 1998: "Invariance-supported photogrammetric triangulation", Proceedings: ISPRS, Commission III, Columbus, OH, pp307-316, July.

Brown, R. G., and Hwang, P. Y. C., 1997: "Introduction to Random Signals and Applied Kalman Filtering", Third Edition, John Wiley & Sons.

Lee, C., 1999: "Mathematical modeling of airborne pushbroom imagery using point and linear features", PhD Thesis, Purdue University, West Lafayette, IN, April.

Mikhail, E. M., 1976: "Observations and Least Squares", University Press of America, New York, NY.

Theiss, H. J., and Mikhail, E. M., 1999: "Georegistration of Tactical Video", Interim Technical Report, Purdue University, West Lafayette, IN, December.