Korean Journal of Geomatics
Vol. 3, No. 1 / June 2003
pp. 23~32

Dynamic Modeling and Georegistration of Airborne Video Sequences

Changno LEE*

Abstract

Rigorous sensor and dynamic modeling techniques are required if spatial information is to be accurately extracted from video imagery. First, a mathematical model for an uncalibrated video camera and a description of a bundle adjustment with added parameters, for purposes of general block triangulation, is presented. This is followed by 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. The results of various experiments with real video imagery, which apply the developed techniques, are given.

Keywords : *Sensor modeling, Video, Invariance, Triangulation*

1. Introduction

Unmanned aircraft vehicles (UAV's) with video cameras on board are becoming a popular reconnaissance technique, particularly for military applications. Major advantages include the speed in data acquisition, and that video cameras and recording media are relatively inexpensive compared to the associated costs of acquiring photography using metric cameras. However, video cameras are generally uncalibrated and inherently contain significant systematic errors that require careful mathematical modeling to describe the path that a ray of light travels from the ground, up through the lens, and onto the sensing medium. Unlike metric cameras, there are no fiducial marks visible on video imagery to allow the precise determination of image coordinates with respect to the principal point, or lens center. So the two principal point offsets and principal distance, or the three basic elements of interior orientation (IO), are generally not well established. A rigorous photogrammetric mathematical model for video imagery must include additional IO parameters to model systematic errors caused by radial and tangential lens distortions, and in-plane distortions caused by non-uniform scaling and skewing of the image pixels.

Rigorous photogrammetric bundle adjustment is based on the collinearity condition equations, which are nonlinear with respect to the unknown parameters. Therefore, the solution is iterative and requires iterations and initial approximations for all parameters. Since video imagery is often captured with oblique views, it is especially difficult to arrive at initial approximations for the exterior orientation (EO) camera parameters. Invariance techniques can be used to obtain camera parameter estimates. Invariance formulations are generally fast and linear, and require no *a priori* information about any of the camera's parameters. While photogrammetry emphasizes the image acquisition parameters, invariance either mathematically eliminates such parameters from consideration or replaces them by non-physically significant parameters that are chosen to make the formulation linear. Consequently, invariance formulations sacrifice accuracy for speed.

For real-time triangulation of video sequences, it is computationally efficient to use a sequential photogrammetry, or Kalman filter, approach. Unlike the batch triangulation approach, which carries all camera parameters from all video frames simultaneously, the sequential approach carries the parameters of the video frame in a state vector that varies with each frame. Although all camera parameters for all frames in the sequence are eventually computed, the sequential

*Senior Researcher, GIS/LBS Project Center, KICT(Korea Institute of Construction Technology (E-mail : changno@kict.re.kr)

The manuscript for this paper was received on April 20, 2003 / accepted on June 30, 2003.

solution inverts matrices of relatively much smaller dimensions. Furthermore, the Kalman filter conveniently incorporates stochastic information regarding the variation of camera parameters between adjacent frames of the video sequence.

2. Photogrammetric Georegistration

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.

2.1 Photogrammetric Mathematical Model

Let the vector $[X\ Y\ Z]^T$ represent the 3D location of any object point, the vector $[x\ y\ z]^T$ represent its image space location, and the vector $[X_L\ Y_L\ Z_L]^T$ represent the location of the exposure station. The expression for the collinearity condition equation in component form is:

$$\begin{bmatrix} x & y & z \end{bmatrix}^T = kM \begin{bmatrix} X - X_L & Y - Y_L & Z - Z_L \end{bmatrix}^T \quad (1)$$

where:

$$x = \bar{x} + \Delta x, \qquad y = \bar{y} + \Delta y, \qquad z = -c \quad (2)$$

and:

$$\Delta x = \frac{\bar{x}}{c}\Delta c + \bar{x}r^2K_1 + \bar{x}r^4K_2 + \bar{x}r^6K_3 + (2\bar{x}^2 + r^2)p_1 + 2p_2\bar{x}\bar{y} + b_1\bar{x} + b_2\bar{y} \quad (3)$$

$$\Delta y = \frac{\bar{y}}{c}\Delta c + \bar{y}r^2K_1 + \bar{y}r^4K_2 + \bar{y}r^6K_3 + 2p_1\bar{x}\bar{y} + (2\bar{y}^2 + r^2)p_2 \quad (4)$$

$$\bar{x} = x_b - x_o, \qquad \bar{y} = y_b - y_o \quad (5)$$

$$r^2 = \bar{x}^2 + \bar{y}^2 \quad (6)$$

in which:

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 three 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).

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, $\sum_{xx}$. 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 a priori covariances.

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 $[x\ y\ 1]^T \approx P[X\ Y\ Z\ 1]^T$ in which (x, y) are image point coordinates, and (X, Y, Z) are ground point coordinates. The "$\approx$" 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 $p34 = 1$.

$$F_{xj} = x_j - \frac{P_1[X_j\ Y_j\ Z_j\ 1]^T}{P_3[X_j\ Y_j\ Z_j\ 1]^T} = 0,$$
$$F_{yj} = y_j - \frac{P_2[X_j\ Y_j\ Z_j\ 1]^T}{P_3[X_j\ Y_j\ Z_j\ 1]^T} = 0 \qquad (7)$$

where j is the point number, $j = 1, 2, \ldots, m$,
P_i is the i^{th} 1×4 row vector of the P matrix,

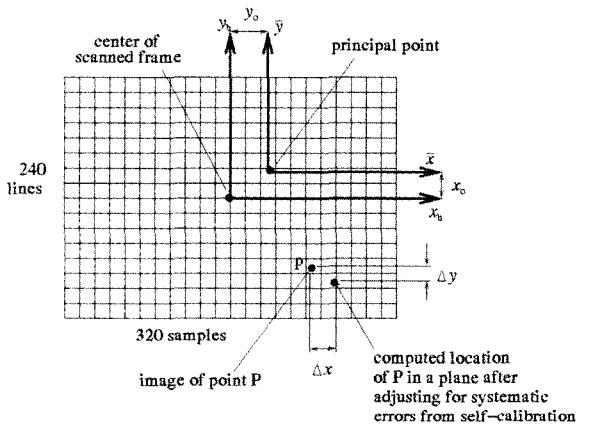

Fig. 1. Video interior orientation.

$i = 1, 2, 3$, and
(x_i, y_i) are the observed image coordinates.

A linear pseudo least squares solution can be applied to Equations (7) 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 (7). Since Equations (7) are nonlinear with respect to the unknowns (p_{ij}), 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):

$$P = k\underset{3\times3}{[AM}\ \vert\ \underset{3\times1}{-AMT}], \qquad T = \begin{bmatrix} X_L \\ Y_L \\ Z_L \end{bmatrix},$$
$$A = \begin{bmatrix} 1 & -b_2 & \frac{-x_o}{c}(1+b_1) \\ 0 & 1+b_1 & \frac{-y_o}{c}(1+b_1) \\ 0 & 0 & \frac{-1}{c}(1+b_1) \end{bmatrix} \qquad (8)$$

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 (8) leads us to arrive at two such constraint equations, G_1 and G_2. If, in addition, 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:

$$G_1 = (R_1R_3^T)(R_2R_3^T) - (R_1R_2^T)(R_3R_3^T) = 0 \quad (9)$$

$$G_2 = (R_1R_1^T)(R_3R_3^T) - (R_1R_3^T)^2 - (R_2R_2^T)(R_3R_3^T) + (R_2R_3^T)^2 = 0 \quad (10)$$

$$G_3 = (R_1R_3^T) - (R_3R_3^T)(x_o^o) = 0 \quad (11)$$

$$G_4 = (R_2R_3^T) - (R_3R_3^T)(y_o^o) = 0 \quad (12)$$

$$G_5 = (R_1R_1^T)(R_3R_3^T) - (R_1R_3^T)^2 - (c^o)^2(R_3R_3^T)^2 = 0 \quad (13)$$

where R = $[R_1\ R_2\ R_3]^T$ = kAM, x_o^o, y_o^o, and c^o are the a priori values of the principal point offsets and principal distance, respectively. Note that any combination of the constraints (Equations 9-13) may be enforced depending on the photogrammetric implications of the problem.

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 a very useful process in engineering applications because its mathematical description is relatively simple and many physical processes are fit well by its stochastic properties. A first order Markov process is a continuous random process that satisfies the criterion that the current state is dependent only on the previous state. Kalman Filtering is a useful technique that allows implementation of sequential least squares adjustment while simultaneously allowing the enforcement of a stochastic process.

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

$$\Delta_i = \Delta_i^- + K_i(f_i - B_i\Delta_i^-),$$
$$K_i = Q_{xx_i}^- B_i^T (B_i Q_{xx_i}^- B_i^T + Q_{e_i})^{-1} \quad (14)$$

where K_i is the Kalman Gain, and Δ_i^- and Q_{xx}^- 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:

$$Q_{xx_i} = (I - K_iB_i)Q_{xx_i}^-, \quad \Delta_{i+1}^- = \phi_i\Delta_i,$$
$$Q_{xx_{i+1}}^- = \phi_i Q_{xx_i}\phi_i^T + Q_{ww} \quad (15)$$

where ϕ_i is the state transition matrix that transforms the current camera parameter state vector of frame i to its predicted state vector of frame (i+1). If we assumed there is no stochastic relationship between different sensor model parameters, 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 com pletion of the forward sweep, the final computations of Equations (14) and (15) result in $\Delta_{n|n}$ and $Q_{xx_{nn}}$, respectively. Proceeding with the backward sweep, updated estimates, the smoothing gain, and the associated covariance matrix can be computed as follows,

$$\Delta_{i|n} = \Delta_{i|i} + K_{s_i}\left(\Delta_{i+1|n} - \Delta_{i+1|i}\right),$$
$$K_{s_i} = Q_{xx_{ii}}\phi^{T}_{i+1|i}Q^{-1}_{xx_{i+1i}},$$
$$Q_{xx_{in}} = Q_{xx_{ii}} + K_{s_i}\left(Q_{xx_{i+1n}} - Q_{xx_{i+1i}}\right)K_{s_i}^{T} \quad (16)$$

5. Experiments

5.1 Photogrammetric Georegistration

Two different data sets with real video imagery are used. The VA Hospital Data set, situated in Leavenworth, Kansas, consists of multiple straight flight lines with a JVC 1085 color video camera. The approximate flying height was 914 m AGL, and the ground sample distance (GSD) at nadir was 3 m. The nominal side-look angle was 30° from nadir. The format size of each video frame was 320 samples and 240 lines, and the pixel size was assumed to be square with 20 μ m sides. The estimated principal distance was 370 pixels. The 320 by 240 pixel format size of a video frame corresponds to a field of view of 56.8°. Control and Check points were extracted on the digital photogrammetric workstation (DPW) from an overlapping pair of triangulated aerial frame photographs. The estimated standard deviations in the X, Y, and Z directions were 0.15 m.

The first experiment is a photogrammetric resection for a single video frame. Frame 41 of strip 7430 was chosen for this experiment (see middle frame of Figure 2).

Several different sets of recovered interior orientation (IO) parameters are tested for comparison. 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.

Each control point forms two collinearity equations. The number of unknown parameters is six for exterior orientation parameter plus the number of IO parameters. Minimum number of control points can be easily computed. For example, case 9 has 10 IO parameters. Therefore, the minimum number of control points is eight.

Several comments can be made after inspecting the results of this experiment. In general, adding IO parameters improves the control point RMS, but not for check points. Recovering a radial lens distortion parameter, K_1, is necessary for resection with an uncalibrated video camera; compare case 1 to case 0.

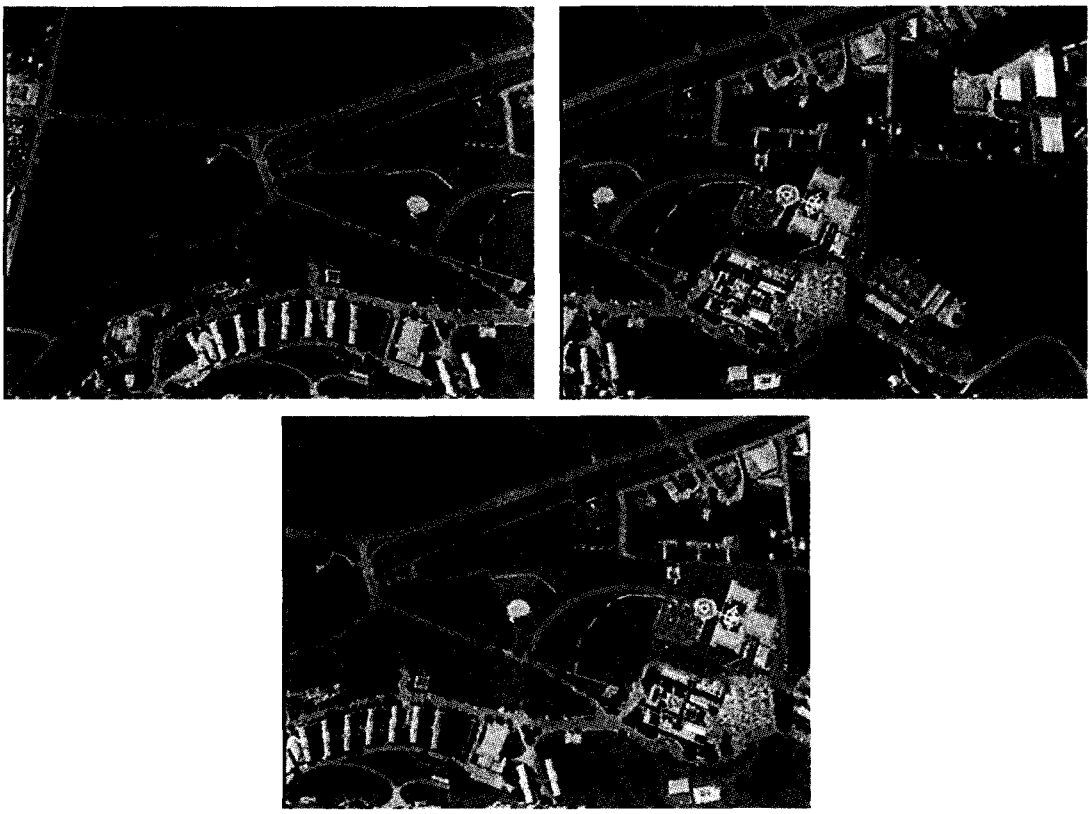

Fig. 2. VA Hospital Data - Frames 7430-1 (top-left), 7430-80 (top-right), and 7430-41 (bottom).

Recovering the principal distance correction parameter, Δc, improves the resection results; compare case 2 to case 1. Estimating the decentering lens distortion and in-plane distortions, (p_1, p_2) and (b_1, b_2), respectively, does not improve check point RMS; compare cases 3-5 to case 2. However, 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. High correlations exist among the camera parameters. For instance, the correlation between the recovered principal distance, c, and the camera location in the height direction, Z_L, has an especially large value of 0.99.

The second video frame georegistration experiment involved the triangulation of three video frames selected from the same strip such that there was approximately 60% forward overlap between successive frames. Resection and intersection are performed simultaneously in the process of the triangulation. Frames 1, 41, and 80 from strip 7430 were used (see Figure 2.). Six cases of recovered IO parameters were tested and the RMS results are tabulated in Table 2. The geometry of the intersection of rays was poor for this experiment because of the relatively large side-looking tilt angle (30° from nadir) and because the pointing directions of all three frames were parallel. Therefore, the values of the Z coordinates were fixed to their known values while the X and Y coordinates of check points were computed as in the single frame case. 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. Although the control point RMS is the lowest for cases 3 and 4, the corresponding check points RMS values are relatively much greater than those for case 2.

The Purdue University data set was taken over the campus in West Lafayette, Indiana. The video imagery was acquired using a VHSC camera on board a Cessna airplane. The camera was manually aimed out of a hole at the bottom of the airplane for the near-vertical and 30? side-looking angle sequences. The flying height for these video frames was 910 meters above the ground. The camera was zoomed out to its smallest principal distance for these sequences, and the estimated principal distance was 740 pixels. The frames in these video sequences were digitized at 640 by 480 pixels; i.e.,

Table 1. Resection Results for Single Uncalibrated Video Frame (33 control; 8 check points)

Case: IO Parameters	Control Point RMS (m)			Check Point RMS (m)		
	X	Y	plan.	X	Y	plan.
0: x_o, y_o, Δc	2.01	2.31	3.06	2.03	1.93	2.80
1: x_o, y_o, K_1	1.38	1.37	1.95	1.41	1.82	2.30
2: x_o, y_o, Δc, K_1	1.28	1.39	1.89	1.31	1.83	2.25
3: x_o, y_o, Δc, K_1, p_1, p_2	1.30	1.33	1.86	1.34	1.89	2.31
4: x_o, y_o, Δc, K_1, b_1, b_2	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	1.30	1.33	1.87	1.33	1.85	2.28
6: x_o, y_o, Δc, K_1, K_2	1.28	1.39	1.89	1.30	1.83	2.24
7: x_o, y_o, Δc, K_1, K_2, K_3	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	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	1.32	1.30	1.85	1.51	2.40	2.83

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

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, Δc, K_1	0.91	0.78	1.20	1.72	2.43	2.97
3: x_o, y_o, Δc, K_1, p_1, p_2	0.88	0.52	1.02	2.83	2.71	3.91
4: x_o, y_o, Δc, K_1, b_1, b_2	0.90	0.48	1.01	2.54	2.67	3.69
6: x_o, y_o, Δc, K_1, K_2	0.90	0.79	1.20	1.72	2.43	2.97
7: x_o, y_o, Δc, K_1, K_2, K_3	0.90	0.80	1.20	1.71	2.43	2.97

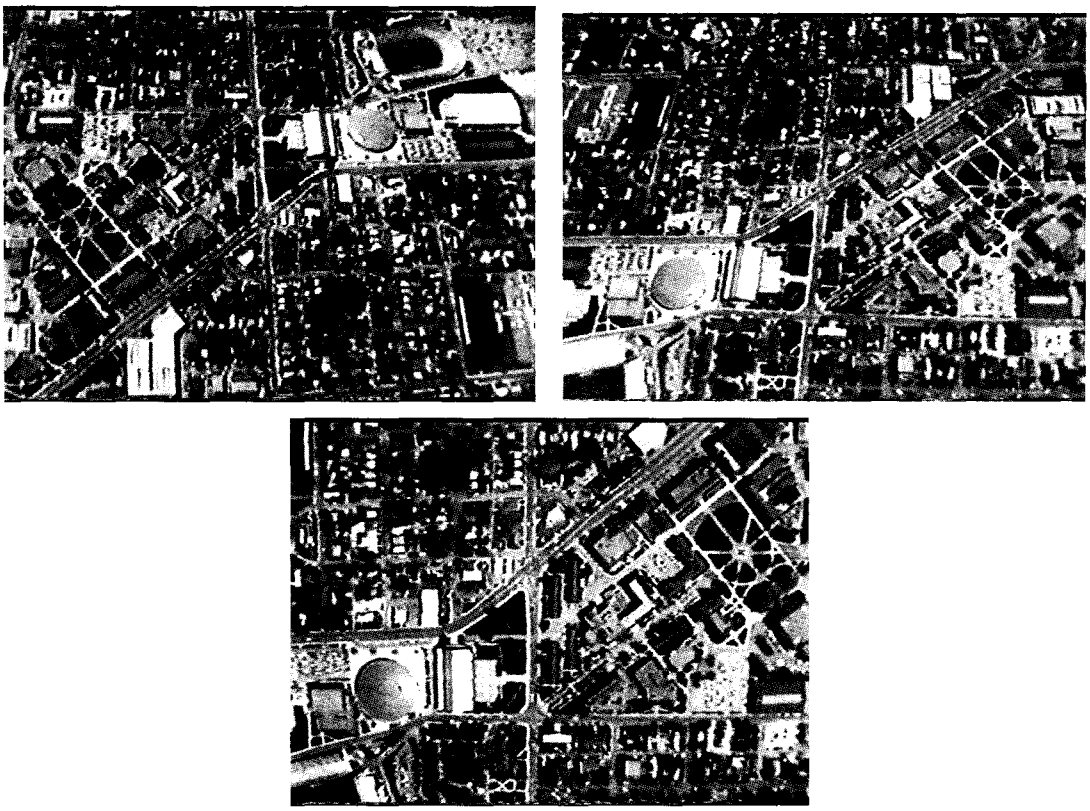

Fig. 3. Purdue Data - Three Convergent Video Frames: Frame 2-158 (top-left), Frame 5-247 (top-right), Frame 3-370 (bottom).

twice the resolution of the VA Hospital data set. Thus, applying this format size and principal distance, the same 56.8° field of view is obtained. A 1:4000 scale block of 27 triangulated near-vertical aerial frame photographs provided plenty of ground control to be used in the following experiments. These 30 micrometer scanned aerial photographs were taken at 580 m AGL, which resulted in a 0.12 m GSD.

A set of 24 points were used as control points in the triangulation of video frames, while 24 points were withheld from the solution and used as check points to evaluate the adjustment. The three convergent video frames, which overlap by nearly 100% are shown in Figure 3. Frames 2-158 and 5-247 each has a side-looking tilt angle of approximately 30°, and all three video frames were extracted from flight lines with a 914 m AGL nominal flying height.

The results for eight cases of recovering different sets of camera parameters are shown in Table 3. The word "Each" for the first two cases in the table indicates that a separate value for each of the listed parameters was recovered for each video frame. The word "All" for cases 3 through 8 indicates that each listed parameter was recovered for the three-frame block collectively. The terms block variant and block invariant are often used to describe the "Each" and "All" cases, respectively. Block invariant parameters are considered to be constant for the whole block. Care must be taken to treat a block as invariant only if the same camera is used with the same zoom and focus settings. Advantages of treating a block as invariant are the higher redundancy and capability of recovering more IO parameters.

The most obvious conclusion to draw from this experiment is the importance of recovering the radial lens distortion coefficient, K_1. In fact, adding it to the typical set of 3 IO parameters reduces the RMS error of ground control and check points by about 25%. Addition of the other IO parameters makes minor but relatively insignificant improvements in some cases.

The relatively high image RMS values can be explained. Recall that the video frames for the VA Hospital data set were scanned at 320×240 pixels while the Purdue data set was scanned with twice as many rows and columns, 640×480. Since the quality of the Purdue video is no better than that of the VA Hospital, the use of a higher resolution was not warranted. Ultimately, the points on the Purdue video

Table 3. Purdue Data - RMS Results for Triangulation of Three Convergent Video Frames, 24 Control and 24 Check Points

Case	Image Check RMS (pixels)			Ground Check RMS (meters)			
	x	y	Radial	X	Y	Z	Radial
Each: x_o, y_o, c	1.78	2.22	2.85	2.22	2.81	2.36	4.29
Each: x_o, y_o, c, K_1	1.11	1.34	1.74	1.08	1.69	2.44	3.16
All: x_o, y_o, c	1.69	2.27	2.83	1.98	2.83	2.67	4.37
All: x_o, y_o, c, K_1	1.06	1.32	1.69	1.01	1.60	2.41	3.07
Each: c All: x_o, y_o, K_1	1.07	1.33	1.71	1.03	1.63	2.42	3.09
All: x_o, y_o, c, K_1, b_1, b_2	1.08	1.33	1.71	1.07	1.61	2.42	3.09
All: x_o, y_o, c, K_1, K_2, b_1, b_2	1.07	1.38	1.75	1.00	1.62	2.34	3.01
All: x_o, y_o, c, K_1, K_2, K_3, b_1, b_2	1.09	1.29	1.69	1.09	1.61	2.31	3.02
All: x_o, y_o, c, K_1, K_2, K_3, p_1, p_2, b_1, b_2	1.09	1.31	1.70	1.08	1.69	2.30	3.05

frames could be measured with accuracy no better than one pixel.

5.2 Invariance Assisted Video Triangulation

5.2.1 VA Hospital Data Results

This experiment was run on Frame 7430-41. Results are shown for four different cases. The first case is for the recovery of 11 parameters with no constraints using linear least squares. Second, results are shown for the case where the scale differential and skew are known to be zeros; i.e., $b_1 = b_2 = 0$. So there are 2 constraints and 9 independent unknowns. The third case shows the recovery of 7 independent unknowns when the principal point offsets are known to be zeros in addition to the differential scale and skew; so there are 4 constraints. The two additional constraints are that $x_o = y_o = 0$. For this third case the seven unknowns include six EO parameters and one parameter for principal distance.

The results from all three cases using the camera transformation matrix and the bundle adjustment case are shown in Table 4. Note that the bundle adjustment results performed significantly better than any of the P matrix based cases. This is due to the recovery of the radial lens distortion parameter, K_1, in addition to the principal distance and principal point offsets.

Table 4. VA Hospital Data - Frame 7430-41: 18 control, 23 check points

No. of Constr.	Check Point RMS (m)		
	X	Y	Planim.
0	2.47	3.00	3.88
2	2.44	3.02	3.88
4	2.41	3.05	3.89
Bundle adj.	1.54	1.93	2.47

5.2.2 Purdue Data Results

This subsection applies the single video frame camera parameter recovery technique to each frame individually for a semi-circular flight line around the EE building on the Purdue campus consisting of 5 frames (Table 5).

Table 5. Purdue Data - RMS Results for Single Frame Parameter Recovery of 5 Video Frames

Frm. No.	No. Ctl, Chk	Check Point RMS (meters)								
		0 constraint			2 constraints			4 constraints		
		X	Y	Plan.	X	Y	Plan.	X	Y	Plan.
3	20, 9	1.59	1.51	**2.19**	1.32	1.61	**2.08**	1.20	1.59	**1.99**
78	21, 10	1.05	1.55	**1.88**	1.05	1.55	**1.88**	0.98	1.56	**1.85**
160	16, 10	1.53	1.26	**1.98**	1.27	1.21	**1.76**	1.17	1.24	**1.71**
233	15, 9	1.84	1.64	**2.47**	0.94	1.27	**1.58**	0.82	1.21	**1.46**
316	18, 9	1.35	2.15	**2.54**	1.10	2.05	**2.33**	1.06	1.96	**2.23**

Note that the number of control and check points is not the same for all frames.

Note that out of the 5 frames, the third case (7 parameters) had the best RMS results. For the cases where some of the IO parameters were not constrained, values were obtained that made no physical sense. When excessively large values are obtained for the IO parameters, the estimated EO parameter values will also be far from their true physical values since there is high correlation among all of the camera parameters, collectively. This camera parameter technique worked reliably for each case with all of the 5 frames, especially since each frame was treated independently with image and ground coordinates as the only input to the algorithm (i.e., no camera information or knowledge about adjacent frames was used).

The camera parameters recovered from the P's were then used as input in a rigorous photogrammetric adjustment.

5.3 Dynamic Modeling

Video sequences from the VA Hospital Data were used to test the developed Kalman filtering algorithm. This experiment consists of 20 non-consecutive video frames from strip 7430. 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.

The ground control points and corresponding image coordinates that were visible on the first frame were used as input to the invariance algorithm to compute camera parameter initial approximations for the first frame only.

Then, approximations for subsequent frames were obtained as part of the Kalman filtering algorithm.

Table 6 shows the check point RMS results for the Kalman filtering Model versus the bundle adjustment for two control configurations: 12 and 6 control points. Initial approximations of the camera parameters for bundle adjustment were computed by the invariance algorithm for all the frames. For this video sequence of frames that spans a relatively long distance, using more control points improves the results, particularly in the Z direction. The Kalman filter performs a little better with 12 control points, while the bundle adjustment performs slightly better with 6 control points.

Another sequence of 20 video frames was selected from a circular trajectory of 180° of arc around the EE building on the Purdue Campus. The first case, with results shown in Table 7, used 11 control points and 11 pass points.

The results for Kalman filtering and bundle adjustment are essentially the same. The reason for similar results is that the selected frames, which are approximately 20 frames apart, are not close enough for a significant stochastic relationship to exist for the dynamic parameters. In this case, the Kalman filter Algorithm essentially becomes a sequential photogrammetric adjustment with results that closely match those from a simultaneous bundle adjustment.

Notice that the Z component of the check point RMS is not the largest component of error, like it typically is for near vertical or slightly oblique cases. The tilt angle for this data set is in excess of 45° to the side from nadir for all frames in the data set. This highly oblique imagery makes for relatively strong geometric ground point intersections in the vertical direction.

Table 6. VA Hospital - 20 Frames from Strip 7430, 20 Pass and 9 Check Points

Number of Control Points	Check Point RMS (meters)							
	Bundle Adjustment				Kalman Filtering			
	X	Y	Z	radial	X	Y	Z	radial
12	2.41	2.21	5.51	6.40	2.55	1.32	5.23	5.97
6	2.93	2.38	7.96	8.81	3.18	1.39	8.32	9.01

Table 7. Purdue Data - RMS Results for Triangulation of 20 Video Frames

Technique	Image Check RMS (pixels)			Ground Check RMS (meters)			
	x	y	Radial	X	Y	Z	Radial
Kalman Filtering	1.88	2.77	3.35	0.87	1.25	0.80	1.72
Bundle Adjustment	1.85	2.78	3.34	0.86	1.25	0.80	1.71

6. Conclusions

The best choice of IO camera parameters appears to be the principal distance, c, the coordinates of the principal point, x_o, y_o, and one coefficient for radial lens distortion, K_1. When attempting to recover more than those four IO parameters, the check point RMS values generally remain the same or increase slightly. Furthermore, solutions that did not include the radial lens distortion coefficient, K_1, result in significantly larger RMS values.

For cases where it is known that the pixels are square (i.e., no skew, and same scale in row and column directions), applying the $b_1 = b_2 = 0$ constraint generally improves the results. Furthermore, knowing that the principal point should be located near the center of the video frame, applying the $x_o = y_o = 0$ constraint generally improves the check point RMS results.

The differences in the results between the simultaneous bundle adjustment (without Gauss-Markov constraints) and the Kalman filtering model (with Gauss-Markov constraints) are generally not significant. Each technique performs slightly better than the other in some of the cases, as indicated by check point RMS results depending on the amount of control. In most cases the RMS of the residuals of the control points used in the bundle adjustment is less than that used for the Kalman filter approach, because the former technique is more flexible due to the absence of stochastic relationship between EO camera parameters.

In this research, covariance between EO and IO parameters are not fully investigated. Instead covariance was checked only when the adjustments didn't converge or showed high check points' RMS errors for a given parameter sets. Therefore it is recommended to analyze covariance for a selection of an appropriate parameter set in the future research.

The results for Kalman filtering and bundle adjustment were essentially the same because the selected frames were approximately 20 frames apart. In the previous experiments, the image coordinates were measured manually. Therefore, it was difficult to use every frame. Automatic extraction of tie points will be investigated for the triangulation of the video sequences in the future research.

References

1. Barakat, H. F., and E. M. Mikhail, 1998. Invariance-supported photogrammetric triangulation, *Proceedings of ISPRS 1998*, Commission III July, Columbus, OH, pp. 307-316.
2. Brown, R.G., and P. Y. C. Hwang, 1997, Introduction to Random Signals and Applied Kalman Filtering, Third Edition, John Wiley & Sons, p. 484.
3. Lee, C., 1999, Mathematical modeling of airborne push-broom imagery using point 4 and linear features, PhD Thesis, April, Purdue University, West Lafayette, IN, p.198.
4. Mikhail, E. M., 1976, Observations and Least Squares, University Press of America, New York, NY, p. 497.
5. Theiss, H. J., and E. M. Mikhail, 1999, Georegistration of Tactical Video, Interim Technical Report, December, Purdue University, West Lafayette, IN, p. 123.