Defence Science Journal, Vol. 52, No. 3, July 2002, pp. 227-242

Implementation of Image Registration Algorithms for Real-time Target Tracking Through Video Sequences

Jharna Majumdar and Y. Dilip

Aeronautical Development Establishment, Bangalore – 560 093

ABSTRACT

Automatic detection and tracking of interesting targets from a sequence of images obtained from a reconnaissance platform is an interesting area of research for defence-related applications. Image registration is the basic step used in target tracking application. The paper briefly reviews some of the image registration algorithms, analyse their performance using a suitable image processing hardware, and selects the most suitable algorithm for a real-time target tracking application using cubic-spline model and spline model Kalman filter for the prediction of an occluded target. The algorithms developed are implemented in a ground-based image exploitation system (GIES) developed at the Aeronautical Development Establishment for unmanned aerial vehicle application, and the results presented for the images obtained during actual flight trial.

Keywords: Unmanned aerial vehicle, image registration, mis-registration, real-time target tracking, target occlusion, cubic-spline model, Kalman filter

1. INTRODUCTION

A study on real-time tracking of targets from a sequence of imagery primarily consists of two major tasks: (i) analysis of image registration algorithms and (ii) applying registration algorithm for tracking applications. In the present scenario, the video images are obtained from an aircraft, which is an unmanned aerial vehicle (UAV) with onboard guidance and navigation system. The aircraft carries a camera in a gimbal which acquires images of the territory and sends the information to a ground control station (GCS) in real-time. During the flight, the pilot in GCS may identify a region of interest, with a mouse click, on the real-time video where a target is present. The target, which appears on a small window, could be tracked by engaging track mode. The position of the target at all the subsequent frames is identified by the window and the information is transmitted to the onboard gyro-stabilisation platform (platform on which the camera is mounted) to correct the azimuth and elevation of the camera so that the target appears at the centre of the frame throughout tracking.

The aim of the authors is to study: (i) the existing image registration algorithms, analyse their computation time, find the accuracy of match points and probability of obtaining false match points, (ii) selection of appropriate hardware for real-time imaging application, (iii) applying registration algorithm suited for target tracking application, and (iv) development of prediction algorithm for occluded target.

2. IMAGE REGISTRATION

Image registration is a process, which finds the location where optimal matching is obtained by matching a template image called the reference

Received 28 February 2002

image, over the searching region of an input image, using suitable similarity measures. The method, although computationally-intensive, is simple, straightforward, and robust, and requires no *a priori* information about the two images.

2.1 Search Strategy

The two *conflicting* requirements of image registration are the time and accuracy. Accordingly, one can employ a fine search (brute force) or a coarse-fine search (efficient on time) method. In fine search strategy, registration starts at the top-left corner of the search space and continues along each row and column, moving the sub-image by one pixel each time. The accuracy of registration algorithm using fine search is good but the computation time is large.

In coarse-fine search strategy, registration is done by extracting sub-images of equal size as that of the reference image, at the start of search space, on a coarse grid at every m^{th} point. An approximate match point is found at the end of this step. A complete search is done in a local region surrounding this match point. The coarse-fine search strategy is efficient if the system allows a minor degradation in accuracy. In order to decide an algorithm suitable for real-time tracking target applications, study of various image registration algorithms available for the selection of an appropriate algorithm is very important.

2.2 Algorithms

A number of image registration algorithms[1-6] were studied on images with varying sizes of search space and reference image. All these algorithms are based on similarity measures between the reference image and the search space. The key issues in image registration are the time required for registration and the accuracy of registration, viz., a measure of how close is the match between the sub-image in the search space and the reference image.

2.2.1 *Histogram-based Approaches*

The basic principle underlying these approaches is to compare the gray-level histogram of the reference image with the sub-image in the search space. Image registration using these methods is less accurate and may give rise to false matches.

2.2.1.1 *Normalised Histogram Intersection Method*

In this method, the sum of minimum gray-level counts for each of the k gray-level normalised wrt reference histogram gray-level results in the match value of k. The best match value is calculated using the formula:

$$\sum_{i=1}^{k} \frac{min\left[Rhist(i), Chist(i)\right]}{Rhist(i)}$$

where $Rhist(i)$ and $Chist(i)$ are the histograms of the reference image and of the sub-image, respectively.

2.2.1.2 *Correlation of Histograms*

In this method, the correlation between the histograms of the reference image and the sub-image in the search space is evaluated. The correlation coefficient value ranges from 0 (for no match) to 1(for perfect match).

2.2.2 *Normalised Area Correlation*

A classical technique for registering a pair of functions is to form a correlation measure between the functions and determine the location of the maximum correlation. A common criticism of the correlation measure form of image registration is the large amount of computation that must be performed if the window and the search regions are large. With this technique, no decision can be made until the entire correlation coefficient values are calculated.

2.2.3 *Statistical Properties-based Approach*

In this method, the statistical properties of the image, viz., brightness, contrast, entropy, and moments are taken as measures for comparison. This method can be used either in stand-alone mode or in conjunction with any of the other methods. The statistical properties for the reference image and sub-image in the search space are calculated and compared. The ratio of standard deviation to the mean (σ_x / m_x) represents the coefficient of variation, and is used for comparison.

2.2.4. *Sequential Similarity Detection Algorithm*

In the sequential similarity detection algorithm (SSDA), the absolute sum of differences between the corresponding pixels in the reference image and the sub-image in the search space is used as the measure of registration. When the registration is perfect, the sum of residual differences has a minimum value but increases rapidly when there is no match. Two SSDA algorithms are: (i) constant-threshold SSDA and (ii) automatic-threshold SSDA.

2.2.4.1 *Constant-threshold* SSDA *Algorithm*

If the sum of residual differences between the best match sub-image and the reference image is known *a priori*, slightly larger value than the sum of residual differences is used as threshold *T*. Execution time decreases as *T* decreases, but if *T* becomes smaller than the minimum value of sum of residual differences, all additions are interrupted halfway.

2.2.4.2 *Automatic-threshold* SSDA *Algorithm*

In this algorithm, addition is first done to completion without threshold value and the resultant sum of residual differences is used as the first threshold value. As the computation progresses, the threshold value is continuously updated by the resultant sum of residual differences obtained for every addition, up to completion. At any time when the computation is in progress, if the new threshold being computed exceeds in value to the one already obtained, the computation can be aborted, thereby saving the computation time.

2.3 Analyses of Sample Images

Analysis of computation time, accuracy of registration, calculation of mis-registration, and effect of reference window size on mis-registration are done for a number of sample images. In all cases, size of the reference image is taken as (50 x 50) whereas the size of the search space is taken as (100 x 100).

2.3.1. *Analysis of Computation Time*

The analysis of computation time for the registration algorithm described above has been carried out for some sample images. It is seen from the timing analysis shown in Table 1 that coarse-fine search strategy using SSDA with automatic threshold has the best performance.

2.3.2. *Accuracy of Registration*

Accuracy of registration is usually calculated using the formula:

$$\frac{\sum_{i=1}^{l}\sum_{j=1}^{m}[R(i,j)-S(i,j)]}{lm}$$

where $R(i, j)$ is the reference image and $S(i, j)$ is the matched sub-image in the search space $l \times m$.

If (x_1, y_1) is the true match point in the search space and the registration algorithm shows a match point at (x_2, y_2), the distance $|(x_2-x_1) + (y_2-y_1)|$ is given a threshold. If this value computed after matching exceeds the threshold, the matching is considered as mis-registration. Table 2 shows the error in matching location using different methods of registration.

2.3.3 *Effect of Reference Window Size on Mis-registration*

To study the effect of size of reference window on the accuracy of registration, SSDA and the area correlation algorithm are run with varying size of the reference window in the search space. Typical results in Table 3 show that the ambiguity in registration starts increasing with the reduction in the size of the reference window in the search space. The ambiguity is due to the fact that multiple points in the search space attain the same match value. From these multiple points, if the farthest point is at a distance greater than a specified threshold, the match point is termed as mis-registration. Mis-registration starts sooner in SSDA than in other methods.

Table 1. Execution time of image registration using various algorithms with fine search and coarse-fine search methods

Images		Algorithm	Time for fine search (s)	Time for coarse-fine search (s)
		Histogram intersection method	18.3610	2.1420
		Cross-correlation of histogram	24.3210	2.6650
		Statistical properties-based	10.2000	1.2143
		Constant-threshold SSDA (SSDA-1)	5.3580	0.6046
		Automatic-threshold SSDA (SSDA-2)	4.3060	0.4791
Search space (100 x 100)	Ref. image (50 x 50)	Area correlation method	36.8230	4.1487
		Histogram intersection method	7.4412	0.8324
		Cross-correlation of histogram	9.9272	1.2120
		Statistical properties-based	4.1450	0.4590
		Constant-threshold SSDA (SSDA-1)	2.1720	0.2419
		Automatic-threshold SSDA (SSDA-2)	1.7451	0.2017
Search space (80 x 80)	Ref. image (40 x 40)	Area correlation method	14.9320	1.7232

3. REAL-TIME TARGET TRACKING

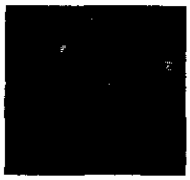
The main issues of real-time target tracking are:

- **Computational complexity of the algorithms:** The algorithms employed for tracking should be efficient on time.
- **Hardware processing of the real-time tracking algorithms:** For a realistic application of target tracking, a suitable image processing hardware capable of acquiring, displaying, and processing in real-time should be chosen.
- **Adaptive tracking:** If the target in the search space varies greatly in size, shape, or orientation from the reference image, no true match can be found and the target in search space may fade into background noise. To be able to track an object while the scene is changing, the reference image must adapt to the changing scene.
- **Walk-off errors during tracking:** If the reference image does not exactly match the target in the search space, the point where the template best matches the target would not correspond to the centre of the target. This produces an error in match value. Although this error is random in nature, it could accumulate from frame-to-frame due to the nature of the tracking algorithm. This error may eventually accumulate to such a point that it walks off from the target.
- **Tracking maintainability in the presence of small interval occlusions:** Target may be partially or fully occluded for some time during tracking. Hence, suitable prediction algorithm should be developed which would be able to predict the target position, in some later time interval, based on the current trajectory.

3.1 Image Processing System for Target Tracking Application

Figure 1 shows the image processing system used for the present application[11]. Real-time image processing system used for target tracking application consists of a pentium PC with a base board for a real-time acquisition and display of frames, and a computation module controller (CMC). Computational module are the commonly used image processing routines implemented in hardware boards. A number of computational modules, plugged-in as

Table 2. Error analysis between the reference image and the search space

Images		Algorithm	Error in matching location
		Normalised histogram intersection method	19.1040
		Cross-correlation of histogram	18.2140
		Statistical properties-based	21.8908
		Automatic-threshold SSDA	10.8160
Search space 100 x 100	Ref. image 50 x 50	Area correlation method	10.8160
		Normalised histogram intersection method	23.6320
		Cross-correlation of histogram	18.3480
		Statistical properties-based	17.6080
		Automatic-threshold SSDA	17.6080
Search space 100 x 100	Ref. image 50 x 50	Area correlation method	17.6080
		Normalised histogram intersection method	13.6090
		Cross-correlation of histogram	11.3612
		Statistical properties-based	12.8180
		Automatic-threshold SSDA	08.9812
Search space 100 x 100	Ref. image 50 x 50	Area correlation method	08.9812

mezzanine modules on each CMC board, accelerate the processing capability of the system. During flight, the system receives two data streams, one is the video data received at a rate of 25 frames/s and the other is the flight data (aircraft and sensor data) received at the serial port of the PC.

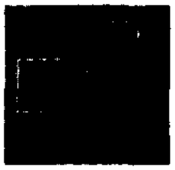
Raw images are acquired by a UAV in an uncontrolled environment where the image quality is affected by dust, smoke, noise, and illumination characteristics of the environment. If the target and the background do not have sufficient contrast, it is required to enhance the contrast using appropriate image preprocessing routines. To achieve the real-time performance, a large number of image pre-processing programs are written using the hardware modules in the CMC and the library calls available in the system. Selection of the preprocessing routines are made adaptive, depending on the nature and the quality of the incoming image[12].

3.2 Target Occlusion While Tracking

Occlusion of a target while tracking may occur through natural or man-made objects in the scene coming into the line of sight of the target or through deliberate deployment of countermeasures, such as smoke. The method of overcoming occlusions is to predict the movement of target using a track memory containing the history of the previous locations of the target.

Two methods for the prediction of target occlusion have been discussed. The first method is based on approximation of the target trajectory using a cubic-spline model[13]. In the present context of target tracking, the target changes its position continuously and the shift in the position (displacement) is not uniform due to some finite-relative velocity of the target. Further, the velocity of the target is not uniform and the rate of change of velocity can be attributed as the acceleration of the target. Based

Table 3. Error analysis for varying window sizes of reference image in a constant search space. The squared blocks in the search space are the mis-registration points

Images	Size of reference image	Algorithm	Error in matching	Accuracy
	70 x 70	Automatic-threshold SSDA	29.189300	Correct match
		Area correlation method	29.189300	Correct match
	50 x 50	Automatic-threshold SSDA	28.478900	Correct match
		Area correlation method	28.478900	Correct match
	40 x 40	Automatic-threshold SSDA	21.419374	Correct match
Search space (100 x 100)		Area correlation method	21.419374	Correct match
	30 x 30	Automatic-threshold SSDA	17.705000	Correct match
		Area correlation method	17.705000	Correct match
Ref. image	20.x 20	Automatic-threshold SSDA	14.975000	Mis-match
(20 x 20)		Area correlation method	14.975000	TRUE
	70 x 70	Automatic-threshold SSDA	18.723000	Correct match
		Area correlation method	18.723000	Correct match
	50 x 50	Automatic-threshold SSDA	17.342000	Correct match
		Area correlation method	17.342000	Correct match
	40 x 40	Automatic-threshold SSDA	12.321000	Correct match
Search space (100 x 100)		Area correlation method	10.787000	Correct match
	30 x 30	Automatic-threshold SSDA	12.321000	Mis-match
		Area correlation method	10.787000	Correct match
Ref. image	20 x 20	Automatic-threshold SSDA	5.257000	Mis-match
(20 x 20)		Area correlation method	5.257000	Mis-match

on the observations about the motion of the UAV and the relative motion of the target, a model that suits the present application, to find the trajectory, is a *cubic-spline model*.

The second method is a spline model Kalman filter[14] (SKF). The filter operates in a predictor-corrector fashion. The accuracy of the prediction using this model depends on the physical modelling of the system and the noise model of the measuring instrument/process. Both these methods work well in case of occlusion.

A check in the match value (min absolute sum in SSDA) triggers the predictive mode. The change in the match value in the case of occlusion is large compared to the signature change due to orientation, and so on. It is also usually much more extensive in the case of occlusion. These two aspects are used to distinguish the two cases.

3.2.1 *Application of Cubic-spline Model*

The mathematical formulation of cubic-spline model is given by Thomson and Green[13]. The observations in the current scenario are X and Y locations of the target in subsequent frames wrt time. To calculate the predicted position (X, Y) of the target at any instant of time, the cubic-spline model has to be applied twice, once for prediction of X with (X, t) as input, and second time for prediction of Y with (Y, t) as input. As soon as the tracking starts, the (X, Y) displacement of the target in every frame is stored in a queue so that at the point of occlusion, the queue has the latest n number of values of (X, Y). Cubic-spline prediction

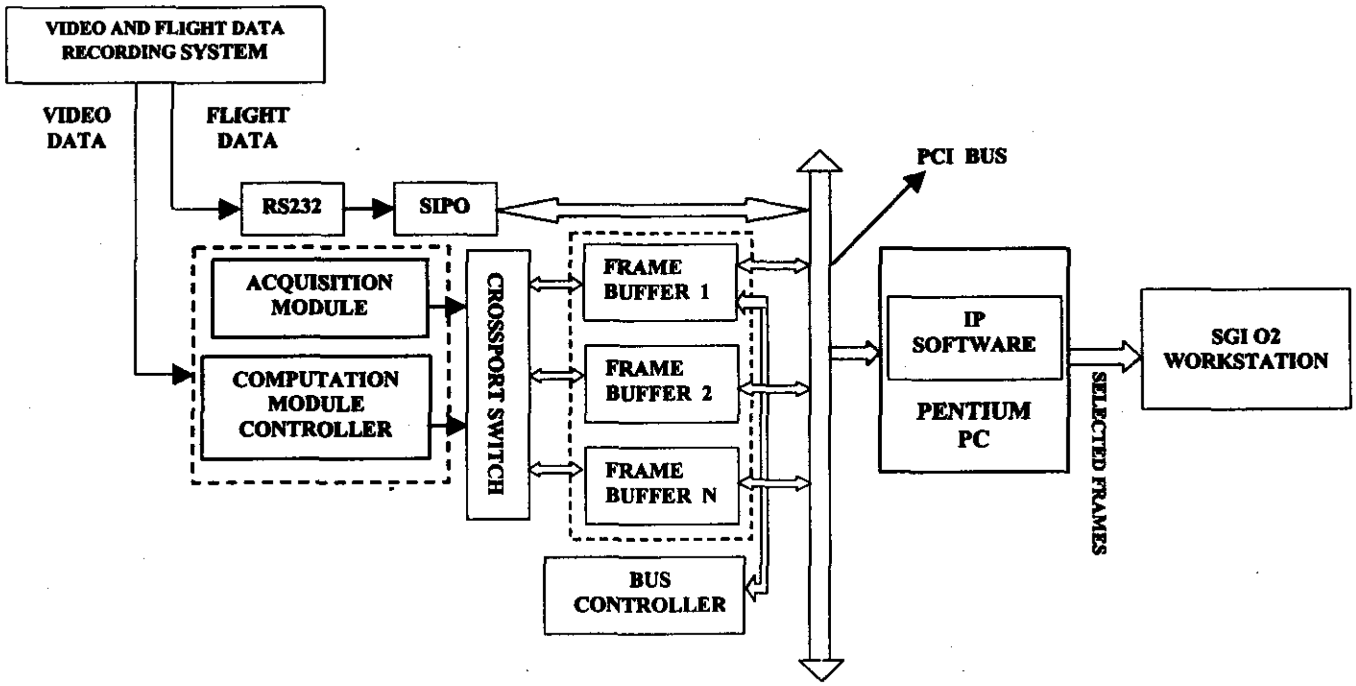


Figure 1. Integrated image Processing station

is now applied to (X, t) and (Y, t) to get the predicted positions of $(Xn+1, t)$ and $(Yn+1, t)$ for the next frame. The newly computed (X, Y) is added to the queue and the prediction is repeated for the next frame. At any instant of time, if the target reappears at the predicted position (determined by searching in the area around the predicted position), the tracking continues and the system locks on to the target.

If the target has an oscillatory motion either in X or Y direction, the cubic-spline model needs to be smoothed. Study has shown that using smoothing and a reduced time step in the computation, the predicted target position during occlusion almost follows the actual observations. The mathematical formulation of smoothing is described in *Appendix* 1.

3.2.2 *Spline Model Kalman Filter*

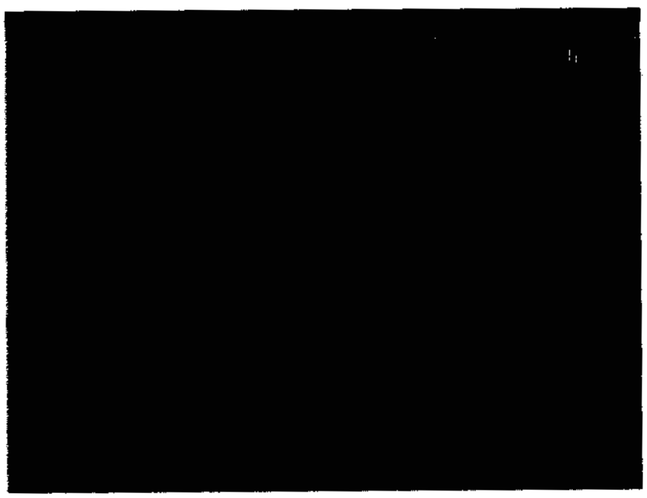
The spline model assumes that the target trajectory can be adequately approximated by a continuous curve, which is twice differentiable. A cubic polynomial satisfies the differentiable property. However, no single cubic polynomial is sufficient to represent a target trajectory in the current scenario of UAV tracking. Applying spline functions, a trajectory can be represented by a cubic polynomial over each of the various time intervals of the trajectory. Physically, an aircraft derives its manoeuvrability from various manoeuvring forces, eg, thrust, lift, and drag which cause changes in acceleration. Rate of change in acceleration is assumed to be constant during a specific time interval. Clearly by specifying the third derivative and duration of time interval, a target trajectory can be approximated to any desired accuracy. The mathematical formulation of spline model Kalman filter that best fits in the current context is available[14].

4. RESULTS & DISCUSSION

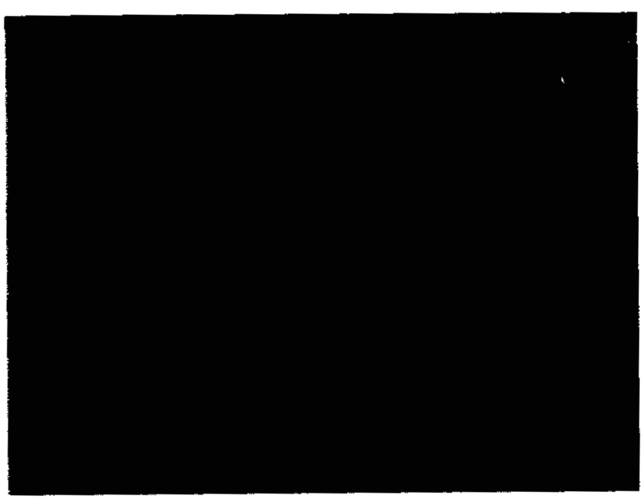
Figure 2 shows the result of target tracking using area correlation method. The tracking sequence is shown at discrete time intervals. The annotated text on these images show the mission time, aircraft, and the sensor parameters required for calculation of geo-location of targets.

Figure 3 shows the result of target tracking using modified SSDA method. Here, raw image without any preprocessing is used for tracking. The target under track is enclosed in a track window of (40x40) pixels and the algorithm employs a

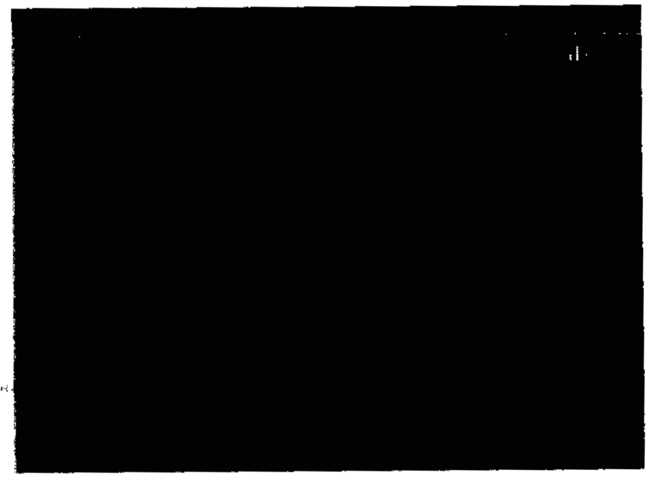
TIME FRAME 1 TRACKING WINDOW



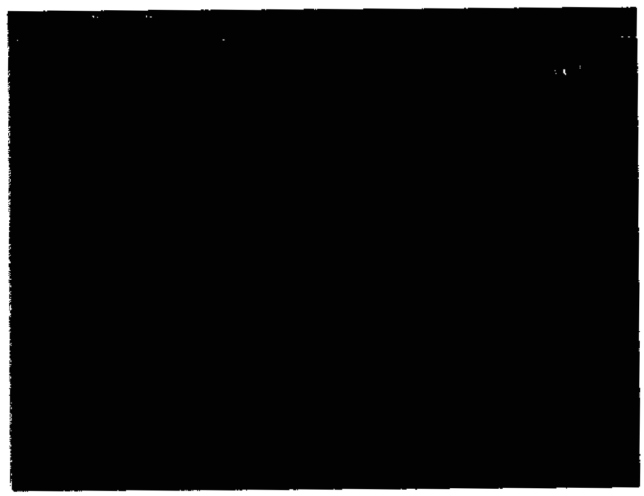
FRAME 2



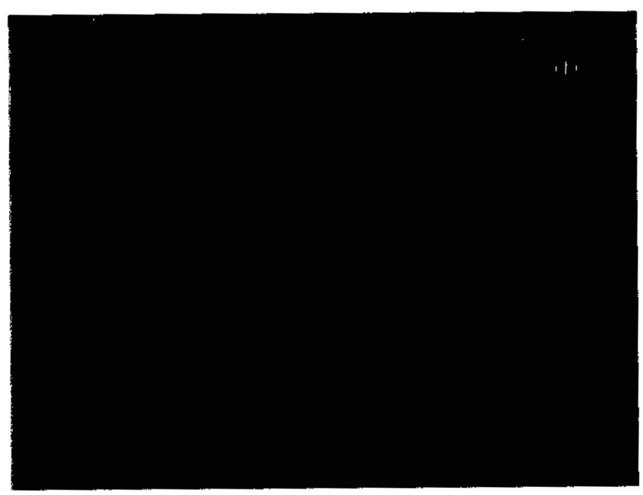
FRAME 3



FRAME 4



FRAME 5



FRAME 6

Figure 2 (a)-(f). Tracking using area correlation method: Tracking sequence: Frame 1 to frame 6: Starting time in frame 1: 15:54:40, ending time in frame 6: 15:55:08.

coarse-fine search strategy (coarse search with step size of 3 pixels) for tracking.

Figures 4(a) and (b) show the (X, Y) location of the target obtained at successive intervals of time during tracking. From the observations it is seen that the target has an oscillatory motion in the Y direction and has a smooth motion in the X direction. This oscillation may occur due to slight disturbance in the motion of the aircraft or due to the vibration of the sensor platform. The analysis shows that the application of cubic-spline model

FRAME 3

FRAME 1

FRAME 4

FRAME 2

FRAME 5

Figure 3 (a)-(e). Tracking using SSDA method. Tracking sequence: Frame 1 to frame 5: Starting time in frame 1: 12:34:40, ending time in frame 5: 12:34:54.

for the prediction of target location is affected by the trend of this oscillation (either decreasing or increasing). This is clear from the results obtained in Fig. 4(c), where the predicted values start increasing if the prediction starts at a point where the oscillatory Y values have an increasing trend at 7.94 s and the predicted values start decreasing if the oscillatory Y values have a decreasing trend at 13.6 s.

Figure 4(d) shows the plot of X, Y values after smoothing. The results in Fig. 4(e) show that the prediction is satisfactory only for two successive time steps. The subsequent predictions deviate from the true value, thereby limiting the prediction while occlusion for just two successive frames/iterations.

Prediction using cubic-spline model is further improved by reducing the time step used for computation from 0.7 s to 0.07 s. With the reduced step size, the predicted positions of the target during occlusion almost follows the actual observations. This is illustrated in Fig. 4(f)

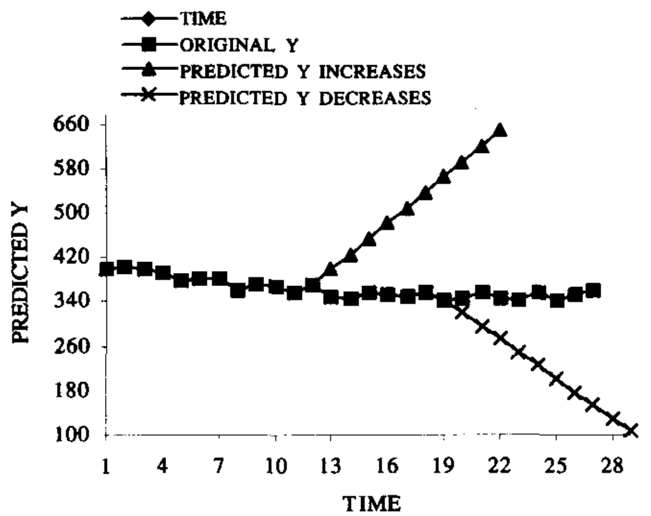


Figure 4(c). Prediction at points where the curvature of *Y* has increasing or decreasing trend

To demonstrate the working of cubic-spline model for prediction while target occlusion, a video sequence is created where a specific target after

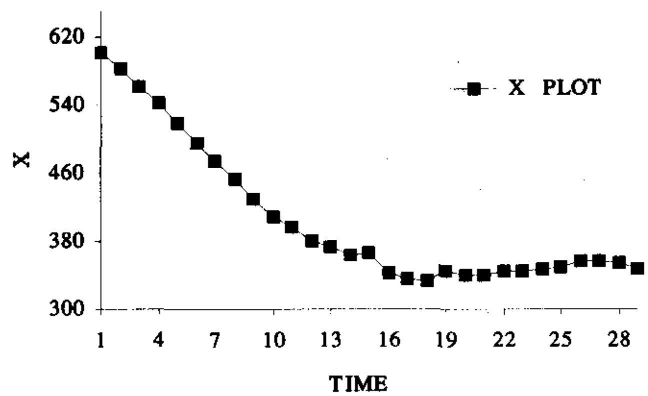


Figure 4 (a). Target trajectory *X* versus time

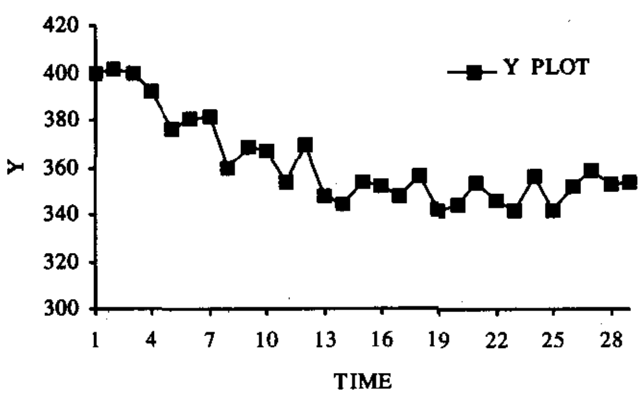


Figure 4 (b). Target trajectory *Y* versus time

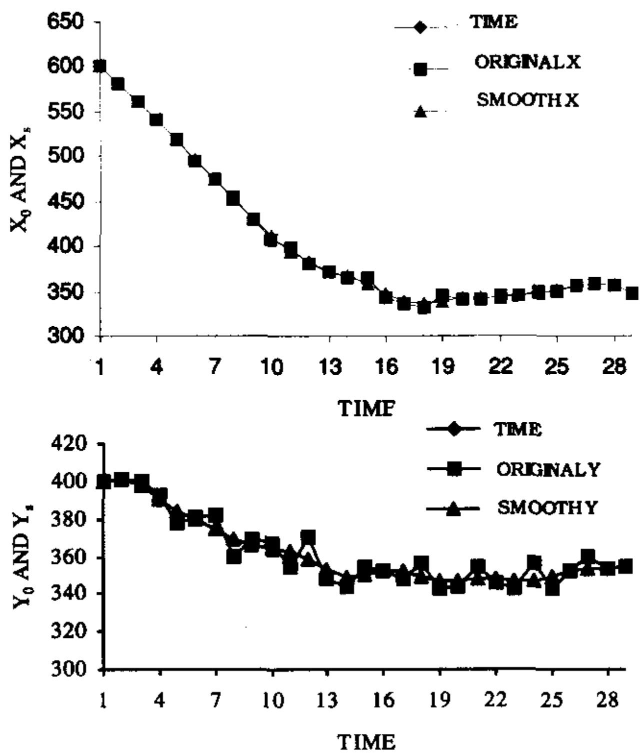


Figure 4(d). Smoothened *X* and *Y* after three Iterations

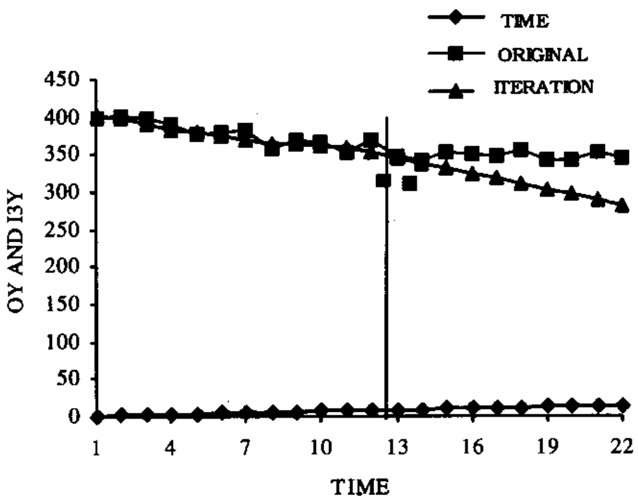


Figure 4 (e). Prediction after smoothing with time step t = 0.7 s.

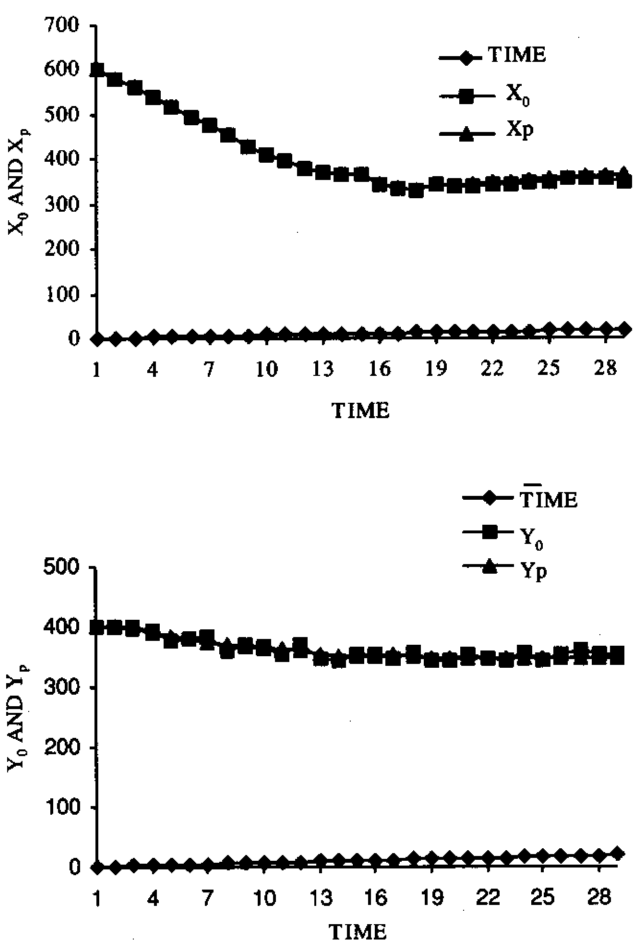


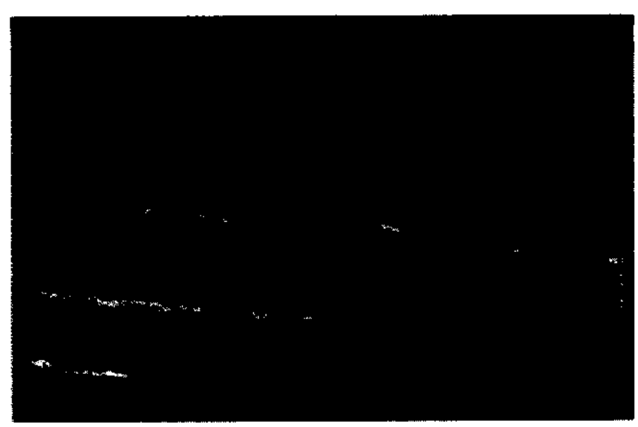
Figure 4(f). Prediction of X and Y with time step t = 0.0712 s.

blanking for 3 s reappears again. The results for 11 frames of successive time instant are shown in Fig. 5. White window in frames 1 to 4 shows the target position while tracking. At frame 4, the target is lost and the prediction starts. Frames 5 to 9 show the predicted position of the target in dotted window using cubic-spline model. Target reappears in frame 10 and tracking resumes as shown by the white window in frames 10 and 11.

The model of the Kalman filter has been applied to the target trajectory (observations obtained by target tracking using SSDA) and simulated using MATLAB tool to ascertain the performance of this model. The initial noise has been assumed to

TRACKING WINDOW

FRAME 1



TRACKING CONTINUES

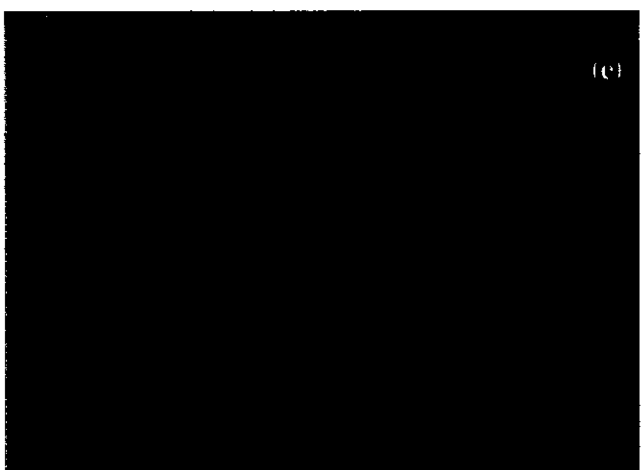
FRAME 2

Figure 5 (a) and (b). Tracking during occlusion using cubic spline method. The tracking sequence is from frame 1 to frame 2.

FRAME 3

At this stage, a flag is given to say that the target is lost. The prediction of target location starts from successive frames.

FRAME 4

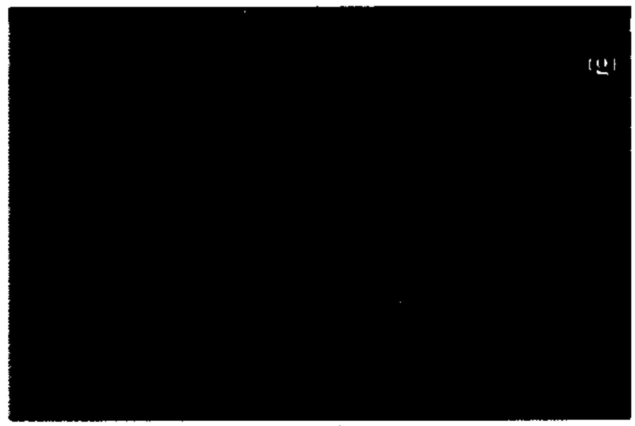


TRACKING WINDOW

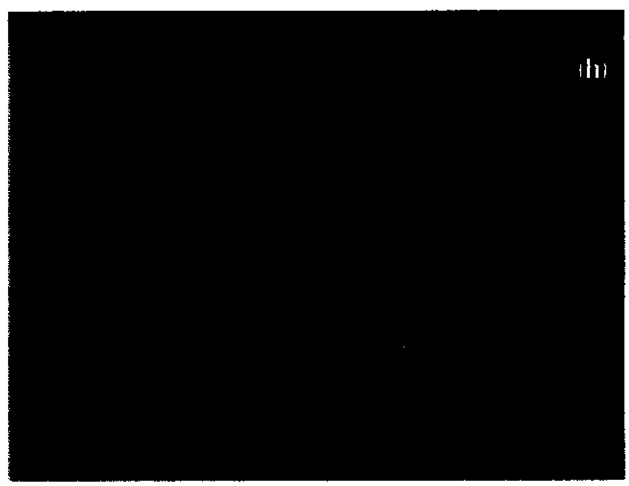
Prediction of target location is shown by the green window.

FRAME 5

FRAME 6

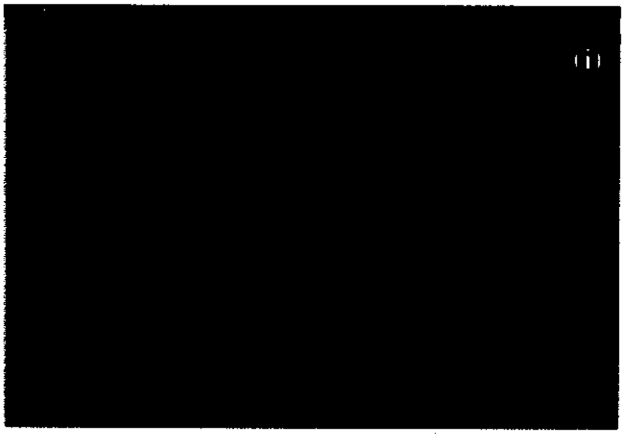


FRAME 7

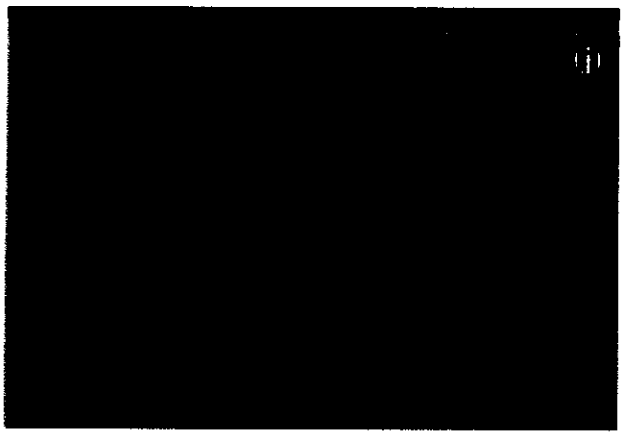


FRAME 8

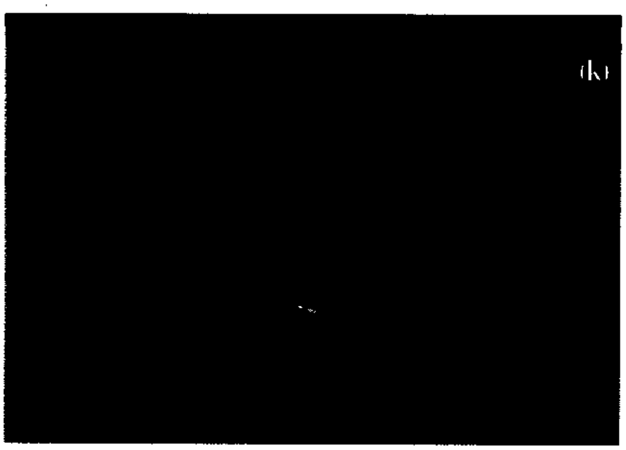
Figure 5(c) to (h). Tracking during occlusion using cubic-spline method. The tracking sequence is from frame 3 to frame 8



FRAME 9



FRAME 10



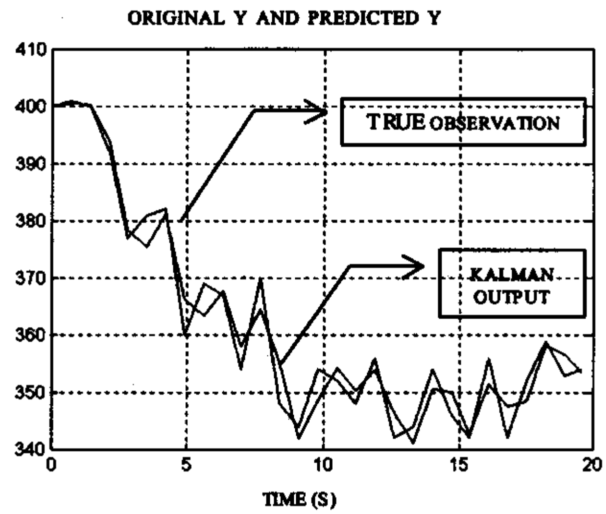
FRAME 11

Figure 5(i) to (k). Tracking during occlusion using cubic spline method. The tracking sequence is from frame 9 to frame 11.

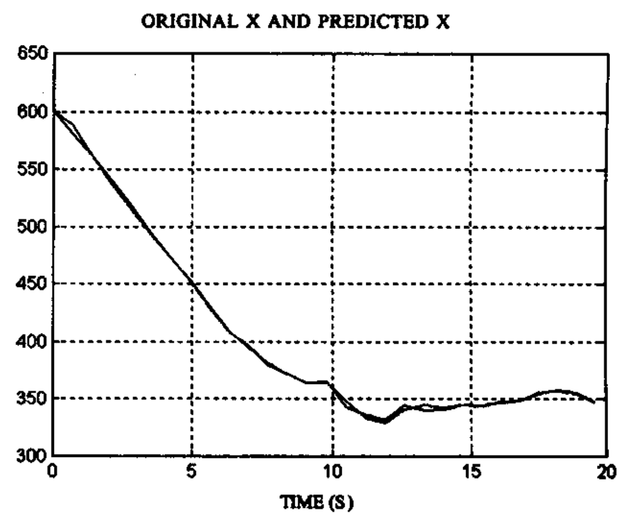
be white noise. The result in Fig. 6 shows that the error is more in the initial phase of prediction as enough observations are not available. But this error (innovation) settles down with time and the model continuously tunes itself iteratively for better predictions. After passage of time, (the point where the innovations settle down), the predicted output of the model matches closely to the true observations.

5. ACKNOWLEDGEMENTS

The authors wish to express their sincere gratitude to Prof Jayanta Mukhopadhyay, Dept of



(a)



(b)

Figure 6. True observation and spline model Kalman filter response: (a) $\{Y, t\}$ and (b) $\{X, t\}$

Computer Science, Indian Institute of Technology, Kharagpur, for valuable discussions during the course of the work. The authors also thank Director Aeronautical Development Establishment, Bangalore for his kind permission to publish this paper.

REFERENCES

1. Prat, W.K. Correlation techniques of image registration. *IEEE Trans Aerospace Electro. Syst.,* 1974, **3**,

2. Barnea, E.I. & Silverman, H.F. A class of algorithms for fast digital image registration. *IEEE Tran. Comput.,* 1972, **C-21**, 179-86.

3. Onoe, M. & Saito, M. Automatic threshold setting for the sequential similarity detection algorithm. *IEEE Tran. Comput.*, 1976, **C-25** (10), 1052-53.

4. Hall, D.J.; Endlich, R. M.; Wolf, D.E. & Brain, A.E. Objective methods for registering land marks and determining cloud motions from satellite data. *IEEE Trans. comput.*,1972, **C-21**, 768-76.

5. Anuta, P.F. Digital registration of multispectral video imagery. *J. Photooptical Instru. Engrs.,* 1969, **7**, 168-75.

6. Svedlow, M.; McGillem, C.D. & Anuta, P.E. Image registration: Similarity measure and preprocessing method comparisons. *IEEE Trans. Aerospace Electro. Syst.,* 1978, **14**(1),

7. Sworder, D.D. & Hutchins, R.G. Image enhanced tracking. *IEEE Trans. Aerospace electro. syst.,* 1989, **25**(5),

8. Das, S.; Mukherjee, J. & Chatterjee, B.N. Analysis of difference pictures for detecting object motion. *In* ICARCV Conference, Singapore, 1990. pp. 866-70.

9. Karim, M. A. Acquisition, pointing and tracking-Special section. *J. Optical Eng,* 1993, 2647-811.

10. Blostein, S. D. & Richardson, H.S. Sequential detection approach to target tracking. *IEEE Trans. Aerospace electro. Syst.,* 1994, **30**(1),

11. Majumdar, J.; Duggal, A. & Narayanan, K.G. Image exploitation for UAV application. *Def. Sci. J.*

12. Majumdar, J. Image processing for remotely piloted aircraft. In National Conference on Applications of Fourier and Wavelet Analysis in Engneering. and Technology (NACWET-99), 1999, Chickbalapur, Karnataka. pp.

13. Thomson, W. E. & Green, R.E. Filtering of high dynamic trajectory data using spline function models. *IEEE Conf. Decision Control,* 1974, 153-58.

14 Gutman, P.O. & Velger, M. Tracking targets using adaptive Kalman filters. *IEEE Trans. Aerospace Electro. Syst.,* 1990, 691-98.