IETE Journal of Research

ISSN: 0377-2063 (Print) 0974-780X (Online) Journal homepage: https://www.tandfonline.com/loi/tijr20

Image Mosaicing from Video Sequences

Jharna Majumdar, B Vanathy, Navneet Varshney, D Raghavendra Rao, Umesh Jalan & Girish R

To cite this article: Jharna Majumdar, B Vanathy, Navneet Varshney, D Raghavendra Rao, Umesh Jalan & Girish R (2002) Image Mosaicing from Video Sequences, IETE Journal of Research, 48:3-4, 303-310, DOI: 10.1080/03772063.2002.11416290

To link to this article: https://doi.org/10.1080/03772063.2002.11416290

IETE Journal of Research
Vol 48, Nos 3&4, May-August, 2002, pp 303-310

Image Mosaicing from Video Sequences

JHARNA MAJUMDAR , B VANATHY, NAVNEET VARSHNEY
Aeronautical Development Establishment, C V Raman Nagar, Bangalore 560 093, India.
D RAGHAVENDRA RAO, UMESH JALAN
Indian Institute of Technology, Kharagpur 721 302,India.
AND
GIRISH R
KREC, Surathkal, Karnataka, India.

The task of assembling individual frames from a video stream into a single giant sized image is known as image mosaicing. The present scenario deals with a real world situation where the video frames and the digital flight data are obtained from a sensor mounted on an aircraft. The paper describes different approaches of automatic image mosaicing suitable for the present application.

The analytical method finds out matching points between the successive image pairs using some sort of similarity measures and use generalized transformation equation to calculate translation, rotation and scaling. The Fourier method uses Fourier transform to calculate the translation and rotation between successive image frames and assemble the images on a large geo-referenced canvas. In both the cases, perspective distortion in the image due to forward-looking sensor is eliminated.

The algorithms, implemented in a SGI workstation, are used to generate mosaics of terrain from the stream of video imagery obtained during actual flight trial of an aircraft and results presented.

Indexing terms : Image mosaicing, Reconnaissance platform, Change detection, Activity tracking, Airborne video surveillance.

IMAGE mosaicing is the task of assembling individual frames in the video stream of data obtained from an aircraft or any other reconnaissance platform. The results of image mosaicing correspond to a simulated wide view representation of a scene. The useful defense application for image mosaicing are in Airborne Video Surveillance for Change Detection, Site Monitoring and Activity Tracking.

The information obtained from mosaicing has a great importance in knowing an unknown terrain. A complete picture of the terrain will help restructuring a large urban area, planning relief in natural disaster like flood or storm. It has also got useful application in document mosaicing, where very high accuracy is required for registering documents.

The basic idea of image mosaicing is that when a video is captured at a sufficiently high rate, there is a significant overlap in the imaged area in the scene between successive images. If this overlap can be detected, it is possible to find out the appropriate geometric transformation required to paste these two contiguous images to obtain an image covering a larger area of the scene. The above process can be continued till all frames in the sequence are merged to generate a single image of very large size [1-6].

PRESENT SCENARIO

In the present scenario [7], camera is mounted on the aircraft, which is a low speed, low altitude aircraft complete with its on-board guidance and navigation system that captures continuous scene of the terrain during flight and sends the information in real time to the ground station. During flight, the system receives two streams of data. The video data is received at the rate of 25 frames/sec in a Pentium PC and the digital flight data, consisting of aircraft and sensor parameters, is received at the serial port of the PC. During flight, frames of interest which require further analysis and interpretation are transferred to a SGI workstation using a high speed network. The image mosaicing software is developed on SGI machine.

METHODS OF IMAGE MOSAICING

The problem of image mosaicing is described in the following steps

- Correction of geometric distortion of each image frame
- Calculation of transformation parameters between successive frames
- Construction of mosaic

Correction of geometric distortion of each image frame

In the present scenario, camera mounted on the aircraft, looks at the terrain at an oblique view angle. This gives rise to perspective distortion in the images.

Calculation of transformation parameters between successive frames

In this paper two different methods are described for calculation of transformation parameters between successive image frames.

- Method using image features
- Method using Fourier Transform

Method using image features [8-9]

The main task of feature based method is accurate and reliable detection of features. Edges, corners, junctions and closed connected regions are usually used as features. Features of one image are used to match with the other image to get the corresponding point. In this section, two different approaches of feature based method are described.

Using generalized transformation equation [10-11]

If two images I_1 and I_2 are translated in x and y directions by (t_x, t_y), rotated by an angle θ and scaled by a factor s with respect to each other, then the rigid planar transformation which map pixel (x_1, y_1) of I_1 to the pixel (x_2, y_2) of I_2 is given by the following equation

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = sR\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + T = s\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad (1)$$

In this method, transformation parameters θ, s and (t_x, t_y) are calculated by matching sub windows of equal size in the overlapped regions of two images using suitable similarity measures. The similarity measure used in the present work is normalized area correlation coefficient, given by the following relation.

$$\sigma = \frac{Sr*SC\sum_{x=1}^{Sr}\sum_{y=1}^{SC} R(x',y')*S(x,y)\sum_{x=1}^{Sr}\sum_{y=1}^{SC} R(x',y')*\sum_{x=1}^{Sr}\sum_{y=1}^{SC} S(x,y)}{\sqrt{\left[Sr*SC\sum_{x=1}^{Sr}\sum_{y=1}^{SC} R(x',y')^2-\left[\sum_{x=1}^{Sr}\sum_{y=1}^{SC} R(x',y')\right]^2\right]*\left[Sr*SC\sum_{x=1}^{Sr}\sum_{y=1}^{SC} S(x,y)^2-\left[\sum_{x=1}^{Sr}\sum_{y=1}^{SC} S(x,y)\right]^2\right]}} \quad (1)$$

Sr and SR are the number of rows and columns of the sub window

$R(x', y')$ is the intensity at (x', y') in the image I_2

$S(x, y)$ is the intensity at (x, y) in the image I_1

A sub window is selected around a feature in the overlapped region in image I_1 and this is matched with the number of sub windows of same size in image I_2 by varying the values of θ and s. At the point of best match, correlation coefficient is maximum and the center point of the sub window gives the translation parameter (t_x, t_y) and the values of θ and s gives the corresponding rotation and scaling parameter.

This method is robust but computationally intensive since it needs exhaustive searching to get accurate result. To reduce the computation time, coarse-fine search strategy is employed.

Calculation of transformation parameters using matching pair of points in two images:

The most critical step to calculate transformation is to find out pair of matching points in two images. In the present method, we have used segmentation approach to find the closed connected regions in the two images, which are then used to find the matching points. The intensity segmentation give rise to number of closely connected regions that are labeled. Geometrical features extracted from these regions are used as matching criterion. The translation, rotation and scaling parameter are obtained from matched pairs by a hierarchical approach.

Let (x_1, y_1), (x_3, y_3) be the first pair of matching points and (x_2, y_2), (x_4, y_4) be the second pair of matching points between images I_1 and I_2. The difference in the slopes of the line joining (x_1, y_1), (x_2, y_2) and (x_3, y_3), (x_4, y_4) gives the angle of rotation θ between the images. After finding the rotation angle θ, the image I_2 is rotated in opposite direction by an angle θ and new coordinates (x_3', y_3') and (x_4', y_4') of (x_3, y_3) and (x_4, y_4) are found. The distance d_1 between the points (x_1, y_1), (x_2, y_2) and distance d_2 between (x_3', y_3'), (x_4', y_4') is calculated and the ratio d_1/d_2 gives the scale factor s. Coordinates x_3', y_3' is divided by the scaling factor to get x_3'' and y_3''. The difference $(x_1 - x_3'')$ gives the translation along x direction and the difference $(y_1 - y_3'')$ gives the translation along y direction.

Method using Fourier Transform [12]

In this approach, Fourier transform is used for calculating the amount of rotation and translation between the images.

If f_1 and f_2 are the two images that differ only by a displacement (x_0, y_0) i.e.,

$$f_2(x, y) = f_1(x - x_0, y - y_0) \quad (2)$$

Then their corresponding Fourier transforms F_1 and F_2 are related by

$$F_2(\xi, \eta) = e^{-j2\pi(\xi x_0 + \eta y_0)} * F_1(\xi, \eta) \quad (3)$$

The cross-power spectrum (i.e., the phase correlation) of the two images f_1 and f_2 with Fourier transforms F_1 and F_2 is defined as

$$\frac{F_1(\xi, \eta) F_2*(\xi, \eta)}{|F_1(\xi, \eta) F_2*(\xi, \eta)|} = e^{-j2\pi(\xi x_0 + \eta y_0)} \quad (4)$$

where **F*** is the complex conjugate of **F**. The shift theorem guarantees that the phase of the cross-power spectrum is equivalent to the phase difference between the images. By taking inverse Fourier transform of the representation in the frequency domain, one obtains a function that is an impulse; that is, it is approximately zero everywhere except at the displacement that is needed to optimally register the two images.

An extension of the translation property can be utilized to determine the amount of rotation as well. If $f_2(x, y)$ is a translated and rotated replica of $f_1(x, y)$ with translation (x_0, y_0) and rotation θ_0 , then

$$f_2(x, y) = f_1(x\cos\theta_0 + y\sin\theta_0 - x_0, -x\sin\theta_0 + y\cos\theta_0 - y_0) \quad (5)$$

According to the Fourier translation property and the Fourier rotation property, transforms of f_1 and f_2 are related by

$$F_2(\xi, \eta) = e^{-j2\pi(\xi x_0 + \eta y_0)} * F_1(\xi\cos\theta_0 + \eta\sin\theta_0, -\xi\sin\theta_0 + \eta\cos\theta_0) \quad (6)$$

If we consider the magnitudes of F_1 and F_2 it is easy to see that the magnitudes of both the spectra are the same, but one is a rotated replica of the other. Rotational movement without translation can be deduced in a similar manner using the phase correlation by representing the rotation as a displacement with polar coordinates, i.e., in polar representation

$$M_1(\rho, \theta) = M_2(\rho, \theta - \theta_0) \quad (7)$$

Above equation is similar to that of eqn (2), Therefore using phase correlation technique we can determine θ_0. Computationally efficient algorithms (FFTs) have been developed for computing the Fourier transform thereby reducing the time complexity of this method. It is theoretically possible to find the scaling parameter using scaling property of Fourier transform. In the present scenario, due to the noise in the images the results were not consistent. We have therefore used a hybrid of Fourier Transforms and analytic technique to find all the three transformation parameters.

Construction of mosaic

After obtaining the transformation parameters, images must be aligned to construct the mosaic i.e., overlaying one image on another. This composing process determines how the pixels in the overlapping region must be represented. In cases where image alignment is close to perfect, it is desirable to use all overlapping images to produce the mosaics. In this case, pixel values in the final mosaic can be computed by averaging the corresponding values of all candidate pixels of image frames that contribute in the generation of the mosaic.

When the alignment between the images is not perfect, average may result in blurring and deterioration of the image quality. In this case it is preferred to choose only one of the input images to represent a region in the mosaic. Such a selection should be done to minimize the effect of misalignment. The most logical selection is to select from each image that part closest to its center since alignment is better at the center than at the edges and distortion at the center of the image is minimum.

Computationally efficient algorithms (FFTs) have been developed for computing the Fourier transform thereby reducing the time complexity of this method. Moreover, the individual frames in a sequence have a good overlap (about 60-70%), which is a prerequisite for this method.

Image mosaicing can be implemented in two different ways: sequential and tree-based. In the first step of sequential image mosaicing, transformation between image 1 and 2 are calculated and used to construct the mosaiced image1-2. In the second step, transformation between image1-2 and 3 are calculated and used to construct the mosaiced image 1-2-3. This process is repeated till the last image i.e., image n is mosaiced with image 1-2-3...$n-1$.

In the tree based image mosaicing approach, for each image pair we construct a sub mosaic. After $n/2$ sub mosaic is completed, where n is the total number of images to be mosaiced, successive sub mosaics are taken and the transformation parameters between them are calculated. This process is continued till one final mosaic is obtained.

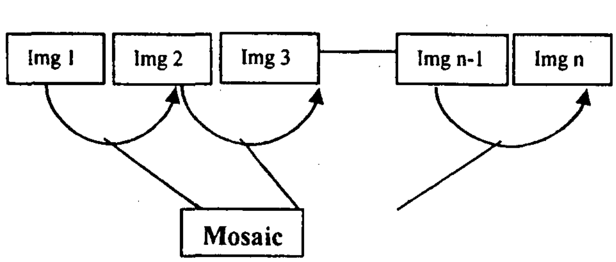


Fig 1*a* The sequential method of mosaicing

Fig 1*a* and *b* show the scheme for sequential and tree based mosaicing approach.

RESULTS AND CONCLUSION

The image mosaicing program is developed on a SGI O2 workstation using C in UNIX environment and the Graphical User Interface is developed using open GL. Both the analytical method and Fourier method are implemented and results presented with the stream of video data acquired during the actual flight trial of an aircraft.

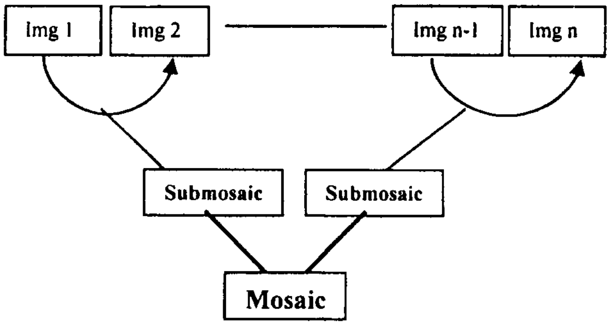


Fig 1*b* Tree based approach of mosaicing

Figure 2 shows an example of mosaicing 8 video frames using analytical method and tree-based approach of merging frames. Figure 3 shows the mosaicing result of 19 video frames using Fourier method and sequential approach of merging frames. Both the approaches have advantages and disadvantages. Figure 4 shows the sequential mosaicing approach using Analytical method implemented on 11 video frames taken from the flight of the aircraft. Figure 5 shows the sequential mosaicing approach using Analytical method implemented on 12 infra red video frames. Figures 6*a* and *b* show the mosaicing of a sequence of video frames using both Fourier and analytical method. The arrow marked area in Fig 6 *a* shows the result of wrong mosaicing. The amount of overlap between two successive frames are very critical in Fourier method which fails to do a correct mosaic in this case since area of overlapped region is too less as shown in Fig 6*c*. Eventhough the individual frames in a sequence have a good overlap (about 60-70%), which is a prerequisite for this method, the actual overlapped region is less after perspective correction. Another example of mosaicing using both Fourier and analytical method is shown in Figs 7*a* and *b*. The marked area in Fig 7*b* shows

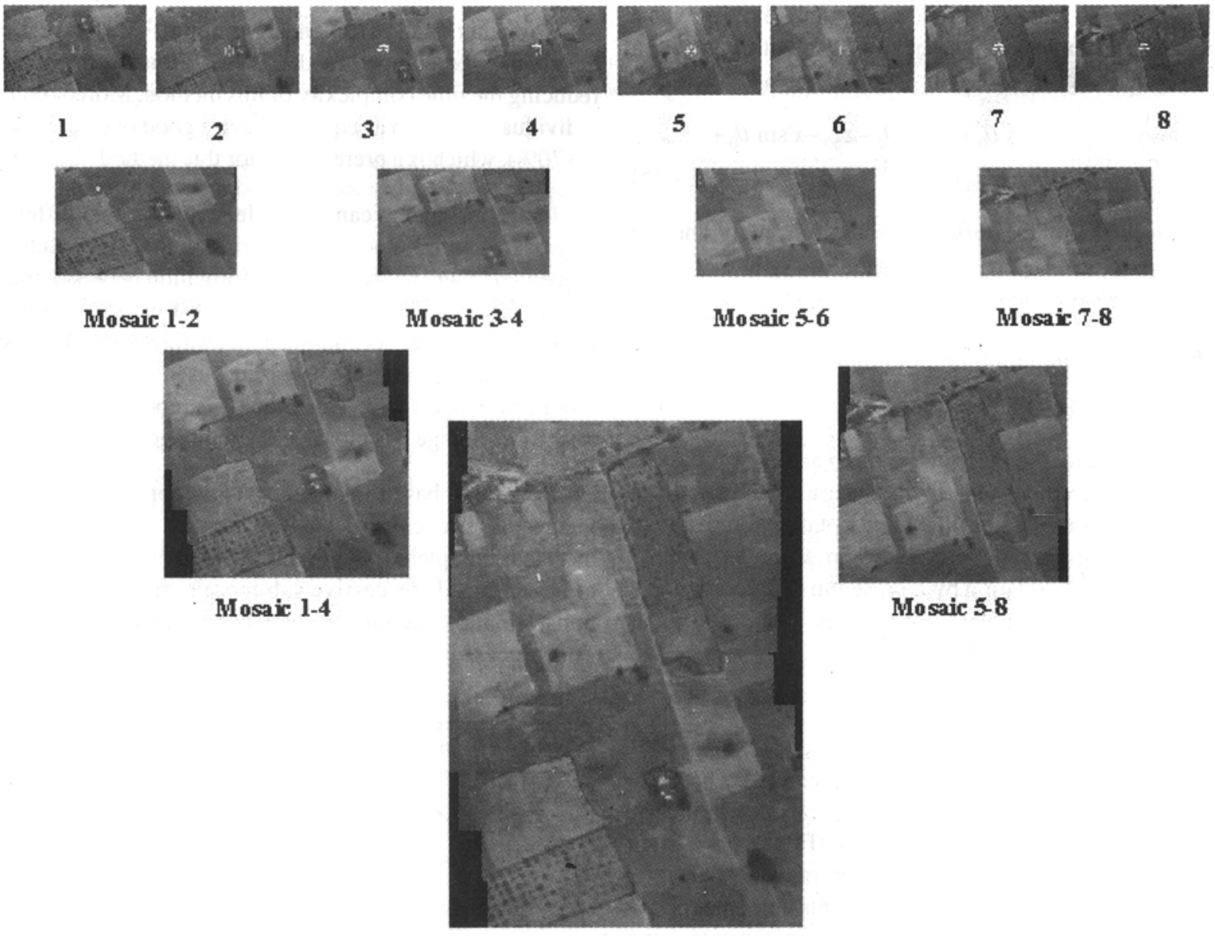


Fig 2 Tree-based mosaicing approach implemented on 8 video frames taken from the actual flight of an aircraft. Mosaic 1-2 is a mosaic of input image 1 and 2 and so on

the result of misalignment in mosaicing using analytical method. The presence of distinctive features in the overlapped region between successive frames is important in analytical method which fails to do a correct mosaic since there is no clear feature as shown in Fig 7*c*.

Automatic mosaicing is done using the video imagery obtained during actual flight trials of a UAV. Both the approaches viz, analytical and Fourier have advantages and disadvantages. The most critical part of mosaicing using analytical method is to find the pair of matching points

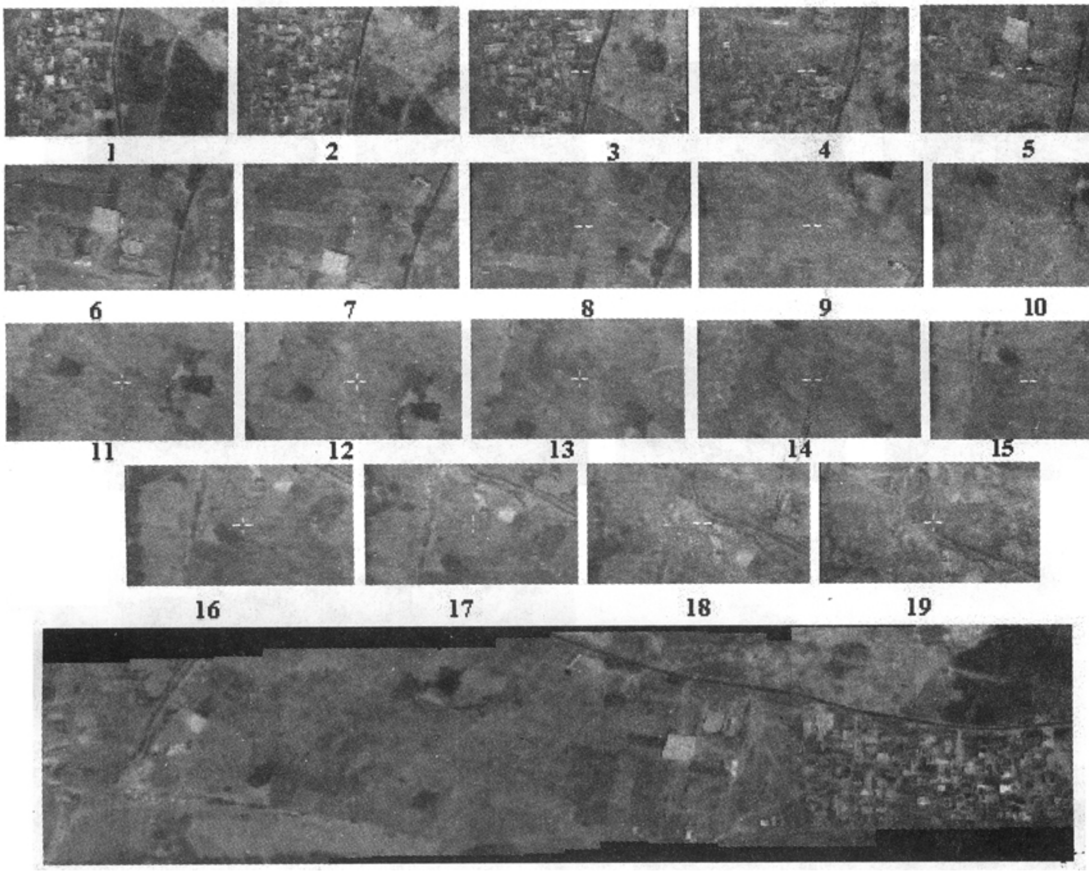


Fig 3 Sequential mosaicing of 19 video frames captured during the actual flight of an aircraft

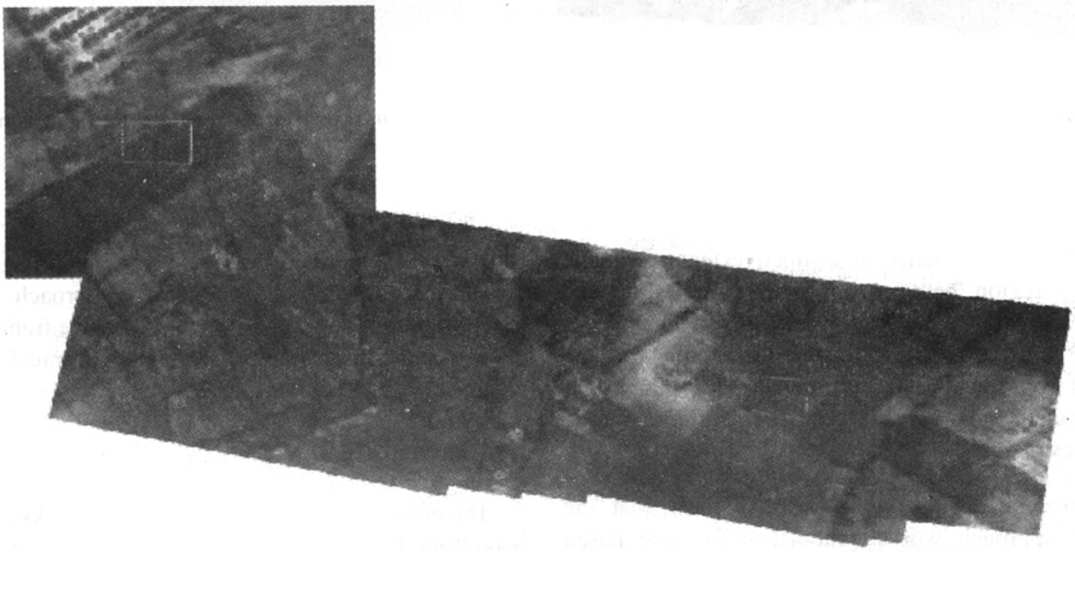


Fig 4 Sequential mosaicing approach using analytical method implemented on 11 video frames taken from actual flight of an aircraft

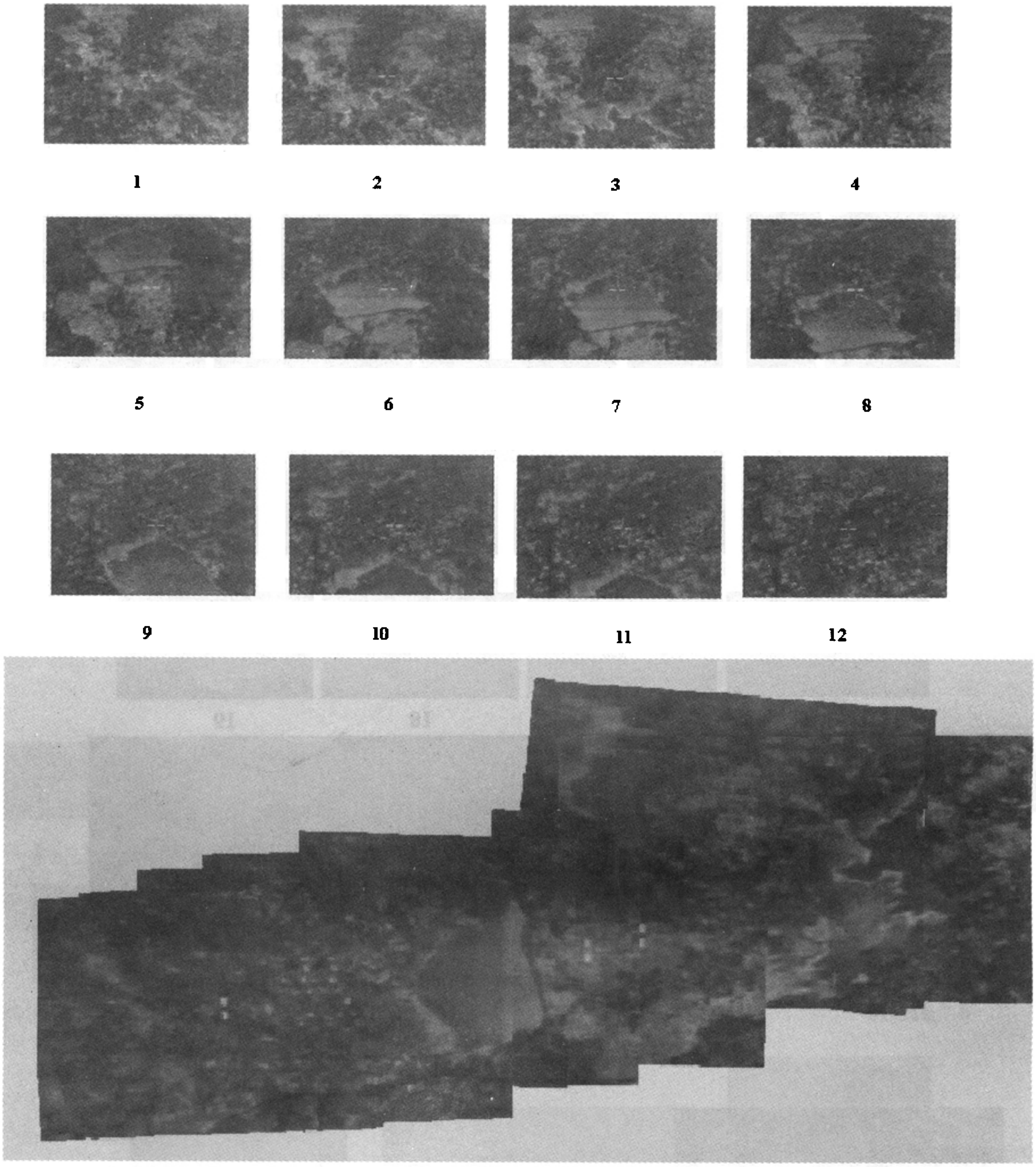


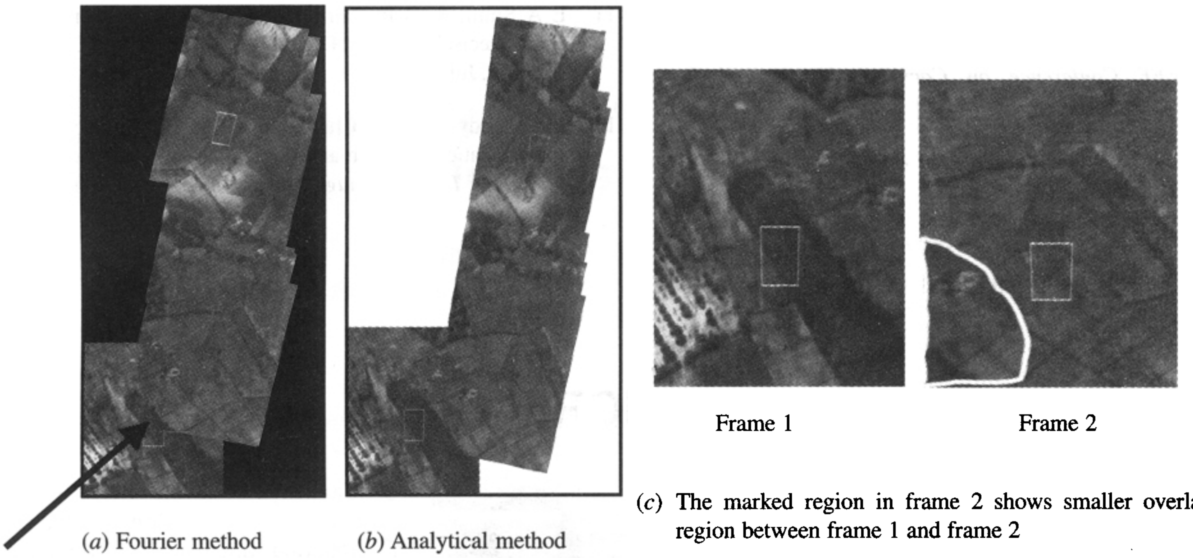
Fig 5 Sequential mosaicing approach using analytical method implemented on 12 infra red video frames from the flight of the aircraft

successfully. The presence of distinctive features in the overlapped region between successive frames is also important in analytical method, which fails to do a correct mosaic if there is no clear feature in the images. The amount of overlap between successive frames is important in Fourier method, which fails to do a correct mosaic if the overlap area is small.

The analysis of the results further show that the sequential approach works better than the tree based approach. This may be attributed to the comparatively large amount of overlap between two successive images to be mosaiced at each stage in the sequential approach. Increase in the amount of overlap between successive frame results in the increase in accuracy of calculation of transformation parameters using Fourier method.

ACKNOWLEDGEMENTS

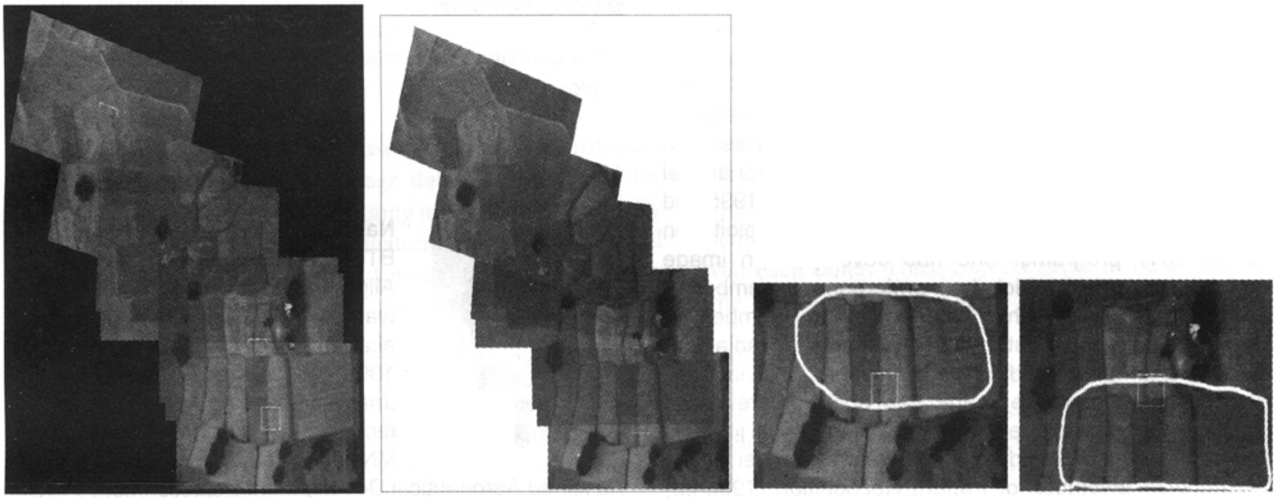
The authors wish to express their heartfelt thanks to Dr K G Narayanan, Director ADE for his constant encouragement, support and kind permission to publish this paper.



Frame 1 Frame 2

(*a*) Fourier method (*b*) Analytical method (*c*) The marked region in frame 2 shows smaller overlapping region between frame 1 and frame 2

Fig 6 Mosaic generated using sequence of video frames from actual flight – the wrong mosaic in (*a*) is shown by the arrow



(*a*) Fourier method (*b*) Analytical method (*c*) The marked region in frame 1 and 2 shows shows the lack of feature in the overlapped region

Fig 7 Mosaic generated using sequence of video frames from actual flight – the wrong mosaic in (*b*) is shown by marked area

REFERENCES

1. L G Brown, A survey of image registration techniques, *ACM Computing Surveys*, 1992.
2. R Szeliski, Video mosaics for virtual environments, *IEEE Computer Graphics and applications*, pp 22-30, 1996.
3. Sevtket Gumustekin. An Introduction to Image Mosaicing. *http://www.pitt.edu/~sevgum/research/mosaicing/index.html*
4. Ismail Haritaoglu, Image Mosaicing. *http://www.umiacs.umd.edu/~hismail/Mosaic/intro.html*
5. Josh Richmond, Image Mosaicing Project, Univ of British Columbia. *http://www.cs.ubc.ca/~jlrichmo/mosaic/index.html*
6. Masakatsu Kourogi, Muraoka Lab Real-time image mosaicing from video sequences. *http://www.murau2.muraoka.info.waseda.ac.jp/members/kourogi/panoV~e.html*
7. Jharna Majumdar, Ankur Duggal, Image Processing for the Remotely Piloted Aircraft, Invited paper presented in the *National Conference on Applications of Fourier and Wavelet analysis in Engineering and Technology*, (NACWET-99), Chickbalapur, Karnataka, May 1999.
8. K Rangarajan, M Shah & D Van Brackle, Optimal corners Detection, *CVGIP*, vol 48, pp 230-245, 1989.

9. I Zoghlami, O Faugeras and R Deriche, Using geometric corners to build a 2d mosaic from a set of images, *In Proc of IEEE Conference on Computer Vision and Pattern Recognition*, pp 420-425, 1997.

10. Anil K Jain, *Fundamentals of Digital Image Processing*, PHI, 1997.

11. E A Smith & D R Philips, Automated Cloud Tracking using Precisely Aligned Digital Pictures, *ibid.*, vol C-21, pp 715-729, July 1972.

12. B S Reddy and B N Chatterjee, An FFT-based technique for translation, rotation and scale-invariant image registration, *IEEE Trans on Image Processing*, vol 5, no 8, 1996.

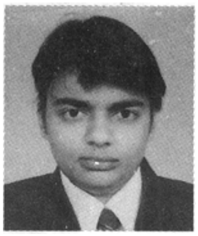
AUTHORS

Jharna Majumdar received her B Tech (Hons) in Electronics and Electrical Engineering and Post Graduate Diploma in Computer Technology from Indian Institute of Technology, Kharagpur in 1969 and 1970 respectively. She received her PhD (Electrical Engineering) in 1980. Dr Majumdar worked as Research Scientist at the Institute of Real Time Computer Systems and Robotics, Karlsruhe, Germany during 1983-1989. She joined DRDO at the Centre for Artificial Intelligence & Robotics (CAIR), Bangalore in 1990 and worked in research areas related to Robot Vision. She joined Aeronautical Development Establishment (ADE), Bangalore in 1995 and since then working in the area of aerial image exploitation for the UAV programs. She has developed an image exploitation system which has undergone a number of actual flight trials. She has guided a number of undergraduate and postgraduate students in India and abroad and presented several papers in national/ international conferences. She has about 50 reviewed technical papers published in national/international journals and conference proceedings. She was a member of the International Institution of Pattern Recognition (1983-89) and IEEE (1982-98) and is a life member of the Aeronautical Society of India.

* * *

B Vanathy received her BE in Electronics and Communication Engineering and ME in Computer Science and Engineering from Manomaniam Sundarnar University in 1997 and 1999 respectively. Ms Vanathy joined Aeronautical Development Establishment (ADE), Bangalore as scientist in the year 2000 and since then is working in the area of aerial image exploitation.

* * *



Navneet Varshney received his BTech in Computer Engineering from Aligarh Muslim University (AMU). He was awarded the Gold Medal for academic excellence by AMU in 1999. His varied interest include Image processing and Broadcasting technologies. He has worked in an MNC for a year before joining DRDO. He joined Aeronautical Development Establishment (ADE), Bangalore as scientist in the year 2000 and since then is working in the area of aerial image exploitation.

* * *