Reprinted from

Symposium on

Machine Processing of

Remotely Sensed Data

June 21 - 23, 1977

The Laboratory for Applications of
Remote Sensing

Purdue University
West Lafayette
Indiana

IEEE Catalog No.
77CH1218-7 MPRSD

MACHINE PROCESSING OF AERIAL DATA FOR AGRICULTURAL RESOURCES INVENTORY AND SURVEY EXPERIMENT

D. S. KAMAT, K. L. MAJUMDER, T. J. MAJUMDAR, I. C. MATIEDA, C.V.S. PRAKASH AND V. L. SWAMINATHAN

Indian Space Research Organisation, SAC, India

I. ABSTRACT

The multiband aerial data for one flight line, stretching over 38 Kms which covers an area of 6528 hectares, of the Agricultural Resources Inventory and Survey Experiment, Patiala has been analysed by an unsupervised automatic processing technique. The results are presented in the form of tables and thematic maps. They are verified with the results obtained independently by visual photointerpretation techniques.

II. INTRODUCTION

The Agricultural Resources Inventory and Survey Experiment (ARISE) project was initiated by Indian Space Research Organisation (ISRO) and Indian Council of Agricultural Research (ICAR) with a view to study soil and landuse and to inventory agricultural produce.[1] For these purposes, two regions namely Anantapur District in Andhra Pradesh, and Patiala District in Punjab were chosen for aerial photography. The basic study was organised around multiband photographs in bands 0.5 - 0.9 micron, 0.7 - 0.9 micron, 0.6 - 0.7 micron and 0.5 - 0.6 micron. In order to collect the photographs, cameras were mounted on an aircraft flying at a height of 7800 ft. The scale of the imagery is 1:30,000.

The Image Processing and Analysis Division (IPAD) in Space Applications Centre (SAC) has used the three spectral bands namely 0.7 - 0.9 micron, 0.6 - 0.7 micron and 0.5 - 0.6 micron for modelling purposes.[2,3] It was found that out of these three band photographs, two bands namely, 0.6 - 0.7 micron and 0.7 - 0.9 micron can give significant results as far as classification of cover types are concerned. There is also a similar view reported in the literature.[4]

In this paper an attempt is made to classify the land cover corresponding to data over Patiala region. ARISE Patiala Campaign was mounted in February - March 1975. Out of the available imagery one complete flight line is selected in a random fashion and the selected flight line, viz. AP7-18W is processed using automatic processing technique. The flight line stretches over 38 Kms and covers an area of 6,528 hectares. The intention to cover one flight line completely was, apart from automatic classification of land cover, to organise for large scale and continuous processing of remotely sensed data using the available computer and other facilities.

One of the needs of work in relation to remotely sensed data and its classification is to extend the methodology from high resolution (scale 1:30,000) imagery like photographs obtained from an aircraft to low resolution imagery (scale 1:1,000,000) obtained from spacecraft borne sensors. As LANDSAT-2 was launched in January 1975, a synchronous LANDSAT-2 imagery over Patiala is available corresponding to dates of ARISE campaign and hence the complete land cover and type classification scheme used for ARISE Patiala data will be helpful in extending the work for LANDSAT type low resolution pictures. In fact the results of automatic processing of ARISE, Patiala data when properly collated with ground truth will provide a dependable strata of information required in analysing satellite imagery.

III. OBJECTIVES OF AUTOMATIC PROCESSING

The objectives of this work are the following :

1. To classify the land cover of Patiala District using automatic and digital classification methods.

2. If the classification method can discriminate between wheat and all other crops then there exists a possibility of extending the procedure to distinguish between wheat and wheat by use of supervised sampling techniques. Wheat is the major crop of Patiala District and it is known that wheat is grown in the district in varieties like W-357, Sonalika, W-RR21 and W-Pb18 and the rest. There is also the need to study the proportions of the various varieties of wheat cultivated in Punjab.

3. There is an interesting fact about the Punjab agricultural practices namely that the areas which are under wheat cultivation and are contiguous generally admeasure 5 hectares or more. Therefore, the automatic processing method used in ARISE Patiala imagery in classifying wheat can be extended to imagery obtained from LANDSAT-2 which has a resolution element admeasuring 70 meters x 70 meters or roughly half a hectare.

4. With the experience obtained from this work procedures can be developed to find a suitable model for processing the multiband imagery using stratified sampling technique.

In this paper, the results in respect of objectives 1 and 2 are presented.

IV. ACQUISITION OF DIGITAL INFORMATION

The photographic imagery in bands 0.6 - 0.7 micron and 0.7 - 0.9 micron is digitized using Optronics Photoscan P-1000 System. This system has the facility to select variable resolutions of pixels in sizes 25 microns, 50 microns and 100 microns and gray scale values lying between 0 to 255 for photographic density range of 0-3D. The imagery for which the results are presented in this paper is digitized using 50 micron resolution. Thus each picture element usually called a pixel, measures 1.5 meter x 1.5 meter on ground. As the actual photographic area on the film imagery is 55 mm x 55 mm there are 1.21 million pixels (bytes) of data in each band. The output from the Photoscan system comes in CCT format wherein one picture element's gray value occupies one byte.

V. DATA STORAGE RETRIEVAL AND PREPROCESSING

A. Data Transfer

The two band digital output in CCTS obtained from Optronics photoscan device is transferred to an input data file consisting of a 29 million bytes capacity IBM 2314 disk for processing purposes. The data storage and retrieval scheme is discussed in detail in a technical note.[5] As the two consecutive photographs are having an overlap of approximately 20%, care is taken that data for two consecutive scenes in both the bands resides in the input data file.

For convenience of handling the data and for other processing purposes, the data in each scene is segmented into a number of strips and each strip into a number of blocks.[2]

B. Data Quality Examination

Once the imagery is digitized and the digital data is transferred to disk, a system of routines are used to examine the quality of data before proceeding to process the data. The system of routines involve routines for gray level printing, frequency distribution, histogram etc.

C. Registration of Multiband Data

In order to classify the two band data it is necessary that data in both the bands for a scene is properly registered. Sequential Similarity Detection Algorithm (SSDA) is used to register the two band imagery.[2,6]

D. Scheme for Analysis

The computer facility available at SAC and Physical Research Laboratory (PRL) Ahmedabad is an IBM-360/44 System and is operated under batch processing mode. Because of this, the processing work has been carried out in different steps. The digitized data for five/six scenes in both bands resides simultaneously in INPUT file. As the analysis progresses, the results of the analysis are stored in TEMP/OUTPUT file. The digitized two band data for the new scene replaces the data of the analysed scene. A block schematic showing the flow of information between various functional units/files is shown in Figure 1.

VI. CLASSIFICATION METHOD

When the information about the number of classes and their signatures either in the form of training fields and test fields or from any other sources are not available for the remotely sensed multispectral data, the obvious approach for analysing the data should be an unsupervised one. In the context of remote sensing one of the unsupervised approaches is clustering i.e. grouping of data on the basis of inherent relationship existing among the elements of each group. Clustering has

emerged as one of the most important techniques for pattern recognition of features or objects of interest, as observed by remotely located sensors.

A. Basis of Classification

The classification method used for the multiband data analysis, as mentioned earlier, is an unsupervised clustering technique and the algorithm used is the Iterative Self Organising Clustering System.[7,2] The basis of the algorithm is as follows: Firstly the data points are assigned to the nearest cluster centre and the cluster centres are modified either by 'combining' or 'splitting' the existing clusters where the nearest cluster centre can be obtained on the basis of the minimum of the distances from the existing clusters. Secondly each class or cluster can be considered as the union of an appropriate number of subclusters of nominal sizes.

B. Procedures used in Classification.

After the multiband data is properly registered and the overlap between two consecutive scenes is removed, the pixels are classified according to the minimum distance criterion.[7] Statistics for each scene, essentially the mean and standard deviation of the clusters, are then computed. The cluster centres are updated each time. The clusters are combined or split according to the specified criterion of maximum standard deviation and intercluster distance respectively.[7] The thresholds for standard deviation, that allow the spreads for classes, and the intercluster distances, which define when two clusters can be combined, are set emphirically.

At the end of the classification, results are stored in the output data file on the disk. The classification information for all the pixels in a scene is stored in a temporary file created on disk and a picture can be reconstructed using that data.

As stated earlier, the statistics for the identified clusters within each scene is stored in the OUTPUT data file. One can assign a gray code value to a picture element belonging to a unique class or a theme. In the final analysis each cover class may be identified with a theme. Using such themes a thematic photomap for the whole scene may be written. However, thematic map writing for a scene is optional. The gray values corresponding to each pixel are stored in a temporary file TEMP from which a thematic photomap is generated.

A program for generating thematic maps reads the data from the TEMP file, formats the data and writes it on magnetic tape which serves as input to the Photo-write system Optronics P-1500 for generating thematic maps on a film transparency.

VII. RESULTS AND DISCUSSION

For the whole area mentioned earlier a total of thirty-six classes are obtained using the classification method discussed in the previous section. This does not account for the classes which constitute much less than 0.1% of the total. Out of the thirty-six classes only twenty-three classes could be assigned a particular class label. These results are presented in Table 1. The remaining thirteen classes could not be assigned any label; the results of which are presented in Table 2. The class labels in Table 1 are assigned on the basis of the interpretation done on false colour infrared imagery by an independent group at SAC.

The false colour infrared transparencies as stated in the introduction earlier, was obtained concurrently with the other multi-band imagery. From Table 1, it can be observed that there are as many as ten variations in wheat and that wheat represents approximately 50% of the identified classes. So many variations in wheat as extracted from the multiband reflectance data may be due to different growth stages, and differences in cultural practices from farm to farm. For some other cover types also like barsim one notices more than one set of spectral reflectance characteristic measure as noted in the Columns 3 and 4 of Table 1.

In order to verify the results obtained by automatic processing of multiband data a scheme was formulated by which it could be possible to check these results against results obtained by visual interpretation. The personnel who did the interpretation were involved in certain field studies which helped them in formulating the interpretation keys.

The picture in Fig.2 is a photograph obtained from a thematic photo generated in a way described in Section VI-B. Figure 3 is a scaled sketch of photointerpretation analysis carried out on the original scene photographed in colour infrared band. Similarity and dissimilarities of the two can be assessed.

In the same fashion thematic photo maps are generated on transparencies for another scene shown in Figure 4 and the corresponding photointerpreted sketch is shown in Figure 5.

One interesting point to be noted in automatic processing technique is that for some covers like human habitation (village residential areas) it was not possible to extract typical reflectance measures which could be used to identify them uniquely. Such areas do form about 6% of land cover in rural areas of Punjab. The automatic interpretation procedure thus accounts for nearly 82% of land cover categories for the areas processed.

The work on automatic processing of multiband photographic imagery relating to a part of Patiala district has given us a certain statistics in respect of well known land covers. The categories of land cover were designated names as available from the results of photo-interpreted data. Using these two sets of information together it was found that over an area of 6528 hectares and during the month of March 1975 the area under wheat was 39% of the total. Other cover types included land under barsim, fallow land, minor crops, sugarcane, bare soil, waste land and formed 29% of the total. There was no characteristic reflectance signature recorded in respect of humand habitation. However, on the recreated thematic photomap it is possible to identify land under human habitation by pecularities of textures etc.

VIII. ACKNOWLEDGMENT

The authors wish to thank Mr.B.S. Joshi for neatly typing the manuscript. Their thanks are also due to the staff of Aerial Surveys, Ground Truth and Photo-interpretation Division, SAC for providing results of interpretation of frames discussed in this paper.

IX. REFERENCES

1. B.Sahai and N.K.Barde, ARISE Tech. Report, Anantapur Campaign, ISRO and ICAR, September, 1975.

2. D.S.Kamat et al, "Use of Aerial Photographs in the Analysis of Landuse", presented at COSPAR XIXTH Meeting, O.M.W.G.Session VI, Philadelphia, PA, USA June 1976.

3. D.S.Kamat, K.L.Majumder and T.K.Nasipuri, "On a Supervised Classification system for Resources Study with Multiband Remotely Sensed Data", Technical Report No.RSA-IP/TN08/76, July, 1976 SAC, Ahmedabad.

4. J.A.Howard, "Concepts of Integrated Satellite Surveys" Third Earth Resources Technology Satellite Symposium Vol. 1, Sec.4, pp 523-531, 1973.

5. "A Programming System for Digital Image Processing at PRL and SAC, Ahmedabad", Technical Report No.RSA-IP/TN09/76, September 1976.

6. D.I.Barnea and H.F.Silverman, "A Class of Algorithm in Fast Digital Registration" IEEE Trans, on Computers Vol.C-21, No.2, February 1972.

7. E.P.Kan et al, "The JSC Clustering Program ISOCLS and its applications" Machine Processing of Remotely Sensed Data, Conference Proceedings, LARS, Purdue University, pp.4B-36, October, 1973.

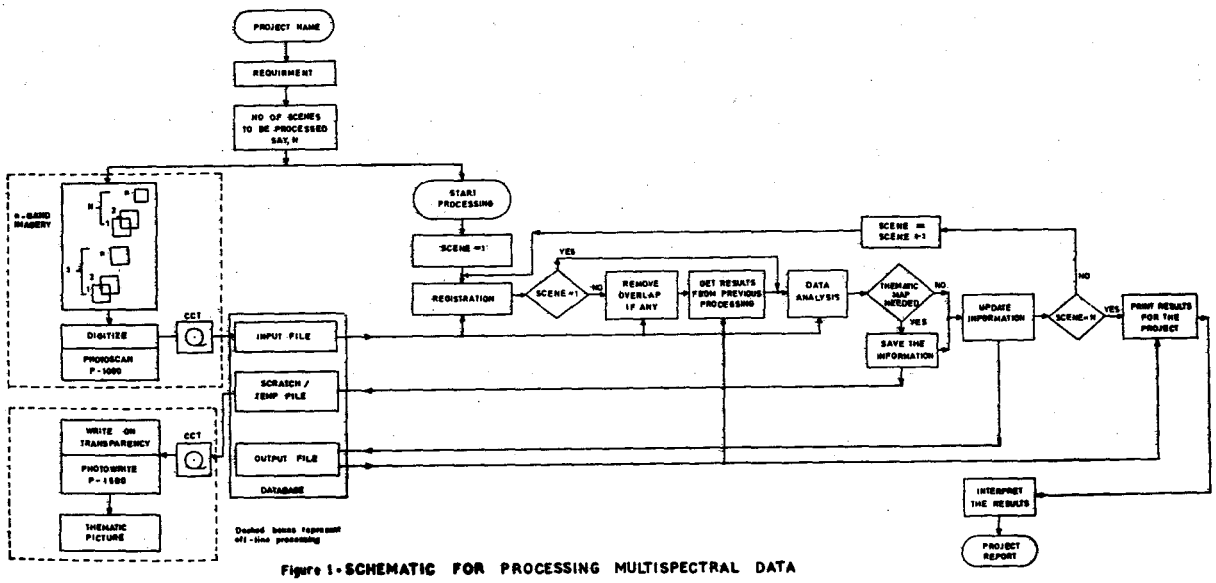

Figure 1.SCHEMATIC FOR PROCESSING MULTISPECTRAL DATA

TABLE-1. STATISTICS OF LABELLED CLASSES.

Class Number	Corresponding ground cover[2]	Mean Gray Value Band 1	Mean Gray Value Band 2	Number of Pixels[1]	Percentage
1.	W	170	175	25,60,753	8.9
2.	W	175	100	9,61,144	3.4
3.	W	196	168	13,88,379	4.8
4.	W	156	148	4,02,006	1.4
5.	W	184	120	16,12,955	5.6
6.	W 80	158	111	5,78,891	2.0
7.	W 18	180	148	17,54,001	6.1
8.	W 80	160	130	8,88,553	3.1
9.	W 357	198	92	4,76,523	1.7
10.	W 21	145	138	1,36,471	0.5
11.	B	186	102	16,31,028	5.7
12.	B	170	180	4,14,332	1.5
13.	FA	162	169	12,48,955	4.4
14.	FA	172	185	7,53,178	2.6
15.	MN	182	160	10,50,080	3.7
16.	SCH	165	137	5,91,257	2.1
17.	SCH	157	158	1,18,741	0.5
18.	SC	142	165	7,85,504	2.8
19.	SC	145	111	92,750	0.4
20.	SC	178	134	9,91,042	3.5
21.	BS	174	186	2,85,293	1.0
22.	BS	182	156	2,08,814	0.8
23.	WL	159	128	78,630	0.3

1) Pixel Size : 1.5 meter x 1.5 meter
2) The interpretation of the scenes were done at AGPD, SAC.
SYMBOLS : - W = Wheat, B = Barsim, FA = Fellow Land, MN = Minor Crops, SCH = Sugarcane Harvested, SC = Sugarcane, BS = Bare Soil, WL = Waste Land.

TABLE-2. STATISTICS OF UNLABELLED CLASSES.

Class Number	Mean Gray Value Band 1	Mean Gray Value Band 2	Number of Pixels[1]	Percentage.
1.	161	82	3,00,314	1.1
2.	172	73	1,99,403	0.7
3.	193	193	7,03,549	2.5
4.	188	212	28,244	0.2
5.	175	197	31,804	1.1
6.	154	197	11,383	0.1
7.	192	144	5,95,759	2.1
8.	203	105	79,395	0.3
9.	213	94	31,429	0.2
10.	126	141	21,236	0.1
11.	137	149	1,15,101	0.4
12.	171	212	33,469	0.2
13.	200	209	29,340	0.1

1) Pixel size 1.5 meter x 1.5 meter.

Figure 2. THEMATIC PHOTO MAP OF FRAME-5

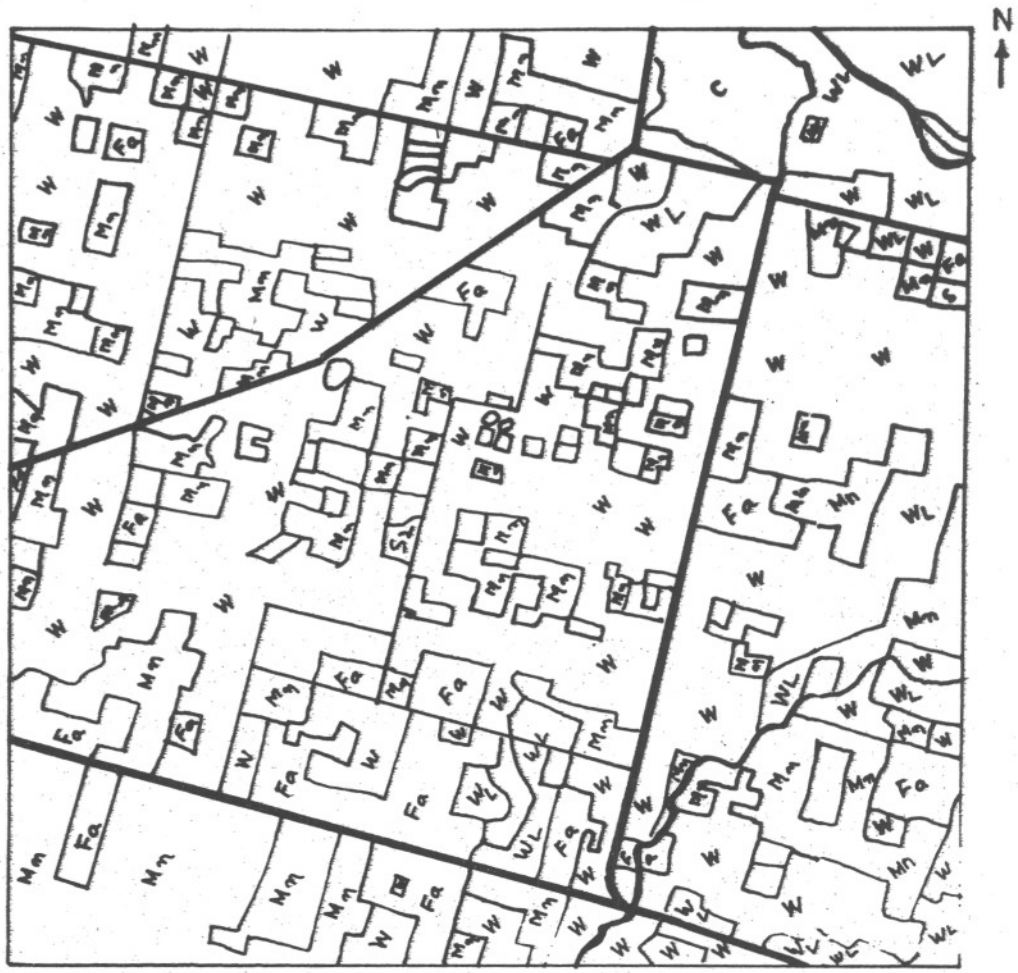

Figure 3 • SKETCH OF PHOTO INTERPRETATION FOR FRAME - 5

Figure 4•THEMATIC PHOTO MAP OF FRAME-26

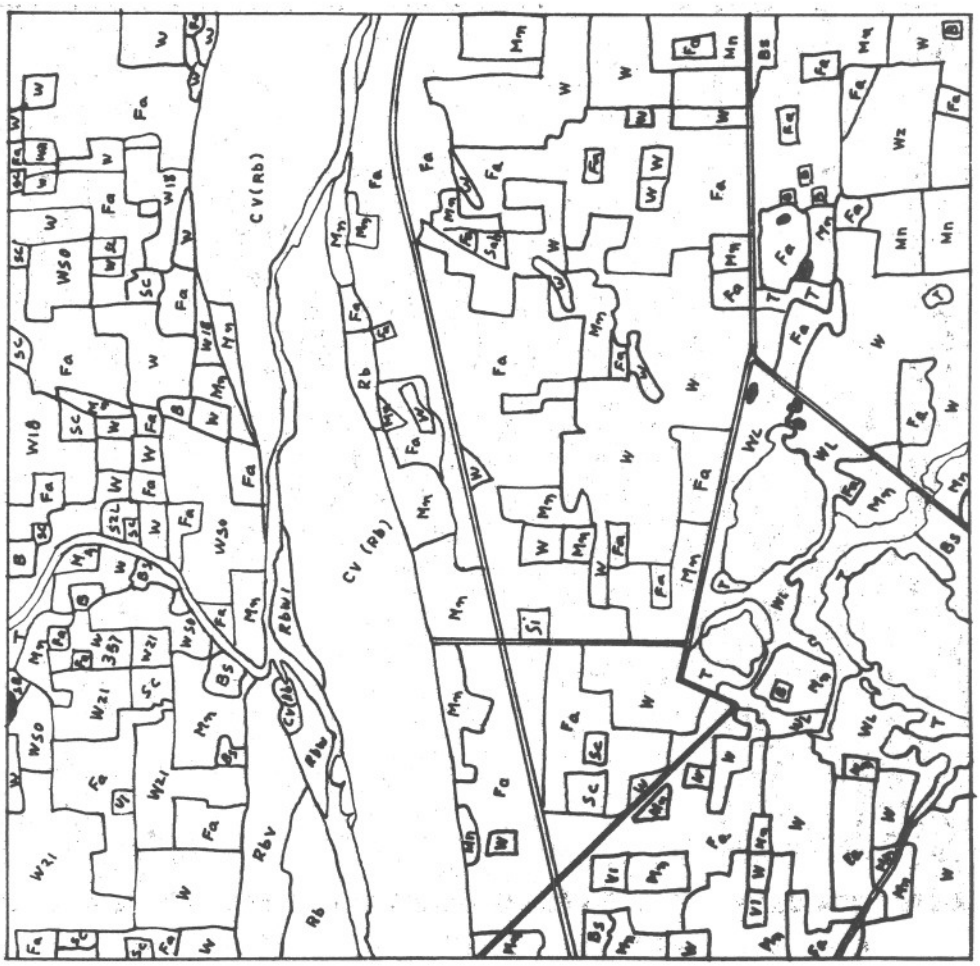

Figure 5. SKETCH OF PHOTO INTERPRETATION FOR FRAME-26

D. S. KAMAT

Studied Physics at University of Bombay, India and later at University of London, United Kingdom; worked in Computer Science and Applications at Indian Statistical Institute, Calcutta, Tata Institute of Fundamental Research, Bombay. Current interest in Remote Sensing for Resource Management.

Mr. T.J.Majumdar graduated in 1969 from Calcutta University with Honours Degree in Physics and then obtained M.Sc. (Physics) degree in 1971 with specialisation in Electronics and Radio Physics from North Bengal University. He was a Research Fellow at the Indian Association for the Cultivation of Science, Calcutta, in Theoretical Physics Division for two years. In July 1976, he joined Space Applications Centre, Indian Space Research Organisation, where he is working at Image Processing and Analysis Division.

Dr. K.L.Majumder got his M.Sc. (Applied Mathematics), Degree in 1968 from Jadavpur University, Calcutta. In 1971, he obtained the diploma in Non-linear Mechanics, DIIT, from IIT, Kharagpur and in 1976 he got his Doctorate degree in Interval Interpolation from IIT, Kharagpur. He joined Space Applications Centre, ISRO in November 1974 and since then working on Image Processing and Analysis of Multispectral remotely sensed data.

Born on 9th October 1951, Mr.I.C. Matieda got his B.Sc. (Mathematics) Degree from M.S.University, Baroda in 1973. He joined Space Applications Centre, Ahmedabad in April 1974. He participated in the ARISE Patiala Campaign, conducted in February-March, 1975.

Dr. C.V.S.Prakash got his M.Sc. Degree in Mathematics from Bangalore in 1971 and Ph.D. from IIT, Bombay in 1975. In April 1975 he joined Space Applications Centre, Ahmedabad, where he has been working in the Image Processing & Analysis Division. Dr. Prakash has been currently deputed to France on a training program.

Mr. V.L.Swaminathan obtained B.Sc. (Physics) Degree in 1962 from Univ. of Madras, and D.M.I.T.(Electronics) Diploma in 1965, from Madras Inst. of Tech., Chromepet, Madras, and M.E.(Electronics) Degree in 1967 from Birla Inst. of Tech. & Science Pilani. From July 1967 to May 1971, he served B.I.T.S.Pilani as a lecturer in Electrical & Electronics Engg. Dept. He joined Electronics Systems Division, ISRO in May 1971 and was actively involved with the Satellite Instructional Television Experiment, using ATS-6 Satellite, conducted by ISRO in collaboration with NASA. Since Aug. 1975, he has been working with Image Processing & Analysis Division, where he is engaged in the development of software packages for analysis of multispectral remotely sensed data.