LARS Technical Report 062685

MULTISOURCE DATA ANALYSIS IN REMOTE SENSING AND GEOGRAPHIC INFORMATION PROCESSING

PHILIP H. SWAIN

Purdue University/Laboratory for the
Applications of Remote Sensing
West Lafayette, Indiana

JOHN A. RICHARDS AND TONG LEE

University of New South Wales
Kensington, AUSTRALIA

ABSTRACT

A general approach is presented for the computer analysis, using quantitative multivariate methods, of remote sensing data combined with other sources of data in geographic information systems. A method is proposed by which inferences can be drawn systematically from multiple observations having significant but unknown interactions. A simple classification experiment with Landsat MSS data is undertaken to illustrate the use of this method.

I. INTRODUCTION

Within the last decade, advances in space and computer technologies have made it possible to amass large collections of data about the surface of the Earth and its environment. More and more typically, these data come from multiple sources: multiple remote sensing systems, digitized terrain information, cadastral data, and so on. Extraction of the great wealth of information contained in such complex geographic data bases requires computer analysis using multivariate quantitative methods. This paper describes a general approach to the development of such methods. Starting from the viewpoint of well known Bayesian classification theory, it explores ways in which inferences can be drawn systematically from multiple observations having significant but unknown interactions and varying degrees of reliability. Emphasis is given to the practical aspects of scene modeling and parameter estimation from available reference data.

The value of exploiting remote sensing data in conjunction with related data from other sources has long been recognized. Data from ground observations are used for classifier training; class prior probabilities are used to minimize overall probability of classifier error; and climatic and meteorological data are used as inputs to crop production estimates. More recently, the availability of digital terrain data has made it possible to utilize topographic information together with remote sensing data for the purpose of land cover analysis.

To a large extent, the methods which have been used for the analysis of multisensor and multisource data have been ad hoc, drawing heavily on the expertise and intuition of the application scientist. Generally applicable methods for assessing and exploiting quantitatively the interactions among the different data sources are not available. The focus of the research reported here is to develop models and analysis techniques, having a sound mathematical/statistical footing, which will facilitate the incorporation into the classification process of as much information as can be determined about multiple data sources and their interactions. Nominally, the approach is through extensions of various modes of pattern recognition; however, any methodology is of interest and may be explored which may serve to implement "convergence of evidence" from multiple sources of information.

II. THE NATURE OF THE PROBLEM

Unlike the situation with purely spectral data in which it is often reasonable to adopt the multivariate Gaussian model, an ensemble of multiple forms of geographic data is bound to exhibit interactions which cannot be prespecified and may be quite complex. To

---

P. H. Swain was on leave during the 1984-85 academic year at the Centre for Remote Sensing and the School of Electrical Engineering and Computer Science, The University of New South Wales, Kensington NSW 2033, Australia.

This research is supported by the National Science Foundation Grant No. ECS 84-00324 and by the Australian Research Grants Scheme.

begin with, the types of data to be combined cannot even be assumed to be commensurable (capable of being expressed in a common units). Example: spectral data combined with elevation data. At minimum, there may be magnitude scaling problems to be dealt with appropriately. In the Gaussian case, variances are used implicitly to scale the data, but this may not be appropriate in other situations. The situation is complicated further in that some data refer to points, some to lines, some to regions; in general these different types will not be commensurable.

Some types of data are inherently non-numerical. Examples: land-use classes, soil types. Although they can be coded numerically for storage in a digital data base, the coding is entirely arbitrary. Such data cannot be treated jointly with other types of data by the more conventional multivariate methods.

The quality of a data source has a bearing on how much influence the source should have on any decision-making process involving a collection of data sources. To date, very little of a systematic or quantitative nature has been done to account for data quality in geographic information processing. In part this may be a result of difficulty in describing or defining data quality. Data quality is manifested in many different ways, such as accuracy, precision, quantization level and reliability. All of these factors should accounted for in models for quantitative analysis of remote sensing and other forms of geographic data.

III. PREVIOUS WORK

A number of very different approaches have been tried for analyzing data from multiple sources. The most straightforward method is simply to form an extended vector with components from all of the data sources and to treat the compound vectors in the same manner as data from a single source. This "stacked vector" approach has been quite successful when the sources are similar and the relations among the variables are easily modeled; e.g., multitemporal data or data from a number of multispectral scanners, analyzed using a multivariate Gaussian classifier [1]. However, this method may not be suitable when the various sources cannot be described by a common model. For example, the multivariate Gaussian model probably should not be used for analyzing extended vectors consisting of spectral data augmented by elevation and slope data. In addition, this approach often involves a substantial penalty in terms of computational cost. When the multivariate Gaussian model is used, the computation time increases as the square of the number of variables.

Other approaches deal with the various sources of data independently. One possibility is to stratify the data based on a subset of sources and then to analyze each stratum based on the remaining sources. In this process, the data are subdivided (stratified) in such a way that variations within each subdivision (stratum) due to some of the variables (the stratifying variables) is minimized or eliminated. For example, this approach has been employed to improve forest cover classification by incorporating information about topography together with Landsat multispectral scanner data [2]. The scene is first stratified into elevation ranges based on available digital topographic data, and then the multispectral data within each stratum are classified into land cover and forest species classes. In carrying out the latter step, prior probabilities are used which are specific to each elevation stratum, thus accounting for the observed relationships between forest species and elevation.

Another alternative is to perform a classification based on one (or more) of the data sources, assess the results, and then resort to other sources to resolve remaining ambiguities. The ambiguity reduction may be carried out by logical sorting methods. For example, Hutchinson [3] describes how slope data were used to resolve the spectral confusion found between the bright surfaces of a dry lake bed and the steep sunny slopes of large sand dunes.

The notion of merging data from multiple sources is explicitly addressed by the method of supervised relaxation labeling described by Richards et al.[4]. In principle, relaxation labeling methods aim to develop semantic consistency among a collection of observations by means of an iterative numerical "diffusion" process. Supervision adds another degree of control of the relaxation process by utilizing an additional source of information. Richards [4] applied the relaxation process to develop spatial consistency in a multispectral classification of mountainous forests, using information about tree species distribution by elevation to supervise the relaxation process. Extensions of this approach to more than two data sources have yet to be investigated. The iterative nature of relaxation labeling makes it computationally very expensive.

IV. THE APPROACH

Although various ad hoc treatments of multisource data have been useful in specific applications, what we are seeking is a general, uniform and widely applicable approach that will capture reliably the information contained in complex data sets while making reasonable demands in terms of the amount of reference data (e.g., ground truth) and computing power required. The method set out here is a first attempt at such an approach. In this section we shall set down the mathematical framework; section VI contains an example illustrating its application.

Let there be n independent sources of data, each providing a measurement x_s, s = 1,2,...,n. Any of the x_s may be measurement vectors. Let there be M information classes (i.e., user-defined classes) denoted w_j, j = 1,2,...,M.

The data from independent sources may be classified into classes most appropriate for the respective sources. These classes are called data classes because they are defined based on relationships in the data space; e.g., spectral classes defined by clustering of spectral data. The ith class from the sth source is denoted by d_{si}, i = 1,2,...,m_s. Measurements are associated with data classes according to a set of data-specific membership functions $f(d_{si}|x_s)$. That is, given a measurement x_s from the sth source, $f(d_{si}|x_s)$ gives the strength of association of x_s with each of the data classes defined for that source.

The concept of data classes is new only to the extent of being a formalization and generalization of the spectral (sub)classes long used in classification of multispectral remote sensing data. Mathematically, the information classes w_j are assumed to be related to the data classes from a single source by means of a set of source-specific membership functions $f(w_j|d_{si}(x_s))$, for all i, j, s. Here $f(w_j|d_{si}(x_s))$ is the strength of association of data class d_{si} with information class w_j, possibly influenced by the value of x_s.

Finally, a set of global membership functions is defined which depends in general on all of the source-specific membership functions. At the global level it will be useful to provide for weighting of the various data sources according to some measure of their "quality," reflecting their reliability or credibility. Thus the membership function F_j for class w_j is of the general form:

$$F_j = F_j[f(w_j|d_{si}(x_s)),\ r_s \mid i=1,2,\ldots,m_s;\ s=1,2,\ldots,n] \tag{1}$$

where r_s is the quality factor for the sth source. A pixel $X = [x_1,x_2,\ldots,x_s]^T$ is then classified according to the usual rule:

Decide X is in class w^* for which

$$F^* = \max_j F_j. \tag{2}$$

The set of global membership functions constitutes a set of discriminant functions for classifying data vectors into information classes.

To implement this very general model the membership functions must be defined specifically. For the present, we shall leave aside consideration of the quality factors, an important matter for future research. Based on Bayesian classification theory, a natural choice for the global membership functions is the posterior probabilities. Let

$$\begin{aligned} F_j(X) &= p(w_j|X) \\ &= p(w_j|x_1,x_2,\ldots,x_s) \end{aligned} \tag{3}$$

Under the assumption that the data sources are statistically independent, this global membership function may be written (see Appendix)

$$F_j(X) = [p(w_j)]^{1-n} \operatorname*{prod}_{s=1,n} p(w_j|x_s) \tag{4}$$

The validity and impact of the independence assumption are discussed further below. Now, each of the source-specific posterior probabilities in the product can be expressed in terms of the data classes. This can be done in many ways. In the following expression,

$$p(w_j|x_s) = \operatorname*{sum}_{s=1,m_s} p(w_j|d_{si},x_s)p(d_{si}|x_s), \tag{5a}$$

the source-specific membership functions appear explicitly as $p(w_j|d_{si},x_s)$ and the data specific membership functions appear as $p(d_{si}|x_s)$. Another useful way to write this may be obtained through straightforward manipulation of the conditional probabilities to get:

$$p(w_j|x_s) = \operatorname*{sum}_{i=1,m_s} p(x_s|d_{si},w_j)p(d_{si}|w_j) \cdot p(w_j)/p(x_s). \tag{5b}$$

Implementation of the classifier for a specific case then involves estimating the various quantities needed to compute equations (4) and (5). The pixel is classified according to (2).

V. THE INDEPENDENCE ASSUMPTION

The assumption of statistical independence used to motivate the product-form global discriminant function, equation (4), deserves further comment. Mathematically, the assumption provides that given two variables x_i and x_j, the joint probability function for the variables is expressible as the product of the marginal probability functions: $p(x_i,x_j) = p(x_i)p(x_j)$. It may be argued that a collection of observations pertaining to a given area on the ground, even though from ostensibly unrelated sources, is unlikely to have the mathematical property of statistical independence. The argument would continue that by adopting such an assumption when it is untrue, one is bound to introduce errors into any decisions based on the associated probabilities. This argument is, of course, well taken. Yet we shall insist on making use of this assumption for compelling reasons.

To begin with, as noted in the Introduction, we are concerned with multiple data sources having complex but unknown interactions. For instance, in order to obtain a regional corn production estimate, there might be available remotely sensed multispectral imagery and soil maps but no explicit reliable information concerning the relationship between vegetation spectral response and soil type. If we are unable or unwilling to collect sufficient ground observations to permit modeling of the soil type/spectral response interactions, our ignorance forces us to treat them as independent variables; we are certainly unwilling to forgo using them altogether. The proposed analysis approach is intended to cope with such a situation.

Another factor is the increased computational complexity which must be accommodated in attempting to deal with the interactions among diverse variables. Even if these interactions are mathematically well characterized, the computational algorithm required to model these interactions may impose a considerable burden on available computer resources.

To the extent that the nature of the dependency among the data variables is known, is believed to be of value in optimizing the analysis results, and can be dealt with within the available computational resources, this information should be utilized using alternative techniques. The approach proposed here provides an avenue for proceeding when these conditions do not hold.

In short, we put forth the relationship between the proposed product-form global membership function and the posterior probabilities as a rationale, not a justification, for the use of the membership function.

VI. AN EXAMPLE

To illustrate the approach set out above, we consider an application requiring the mapping of forest species in an area of rugged terrain. It has been demonstrated that analysis of multispectral data augmented by elevation data can produce better forest species classification than can analysis of multispectral data alone [1,2,4].

Let $X = [x_s, x_e]^T$, where x_s is a vector of spectral measurements and x_e is elevation.

First we attend to the spectral data. The data classes corresponding to x_s are spectral classes which may be derived by any of the usual supervised or unsupervised classifier training methods. If clustering were used to define the spectral data classes by unsupervised classification, the spectral classes d_{si} might then be defined by the maximum likelihood rule

x_s is in d_{si} iff

$$p(x_s|d_{si}) = \max_j \, p(x_s|d_{sj}). \qquad (6)$$

Equation (5b) above can be used to compute the posterior probabilities associated with the spectral data. Each of the conditional probabilities $p(x_s|d_{si},w_j)$ may be modeled by a multivariate normal density function with parameters estimated from the training sample and the clustering results; each conditional probability $p(d_{si}|w_j)$ may be estimated by the fraction of the training sample for class w_j classified into spectral class d_{si}; and the prior probabilities $p(w_j)$ may be estimated in the usual way, such as from a representative training set.

The elevation data classes, corresponding to x_e, are simply elevation ranges. The posterior probabilities $p(w_j|x_e)$ must be estimated from information about the distribution of tree species as a function of elevation (see [1,2]).

Thus the set of global membership functions for this problem, based on equation (4), is

$$F_j(X) = [p(w_j)]^{-1} \, p(w_j|x_s) \, p(w_j|x_e) \qquad (7)$$

$$= \sum_{i=1,m_s} p(x_s|d_{si},w_j)p(d_{si}|w_j)p(w_j)/p(x_s) \cdot p(w_j|x_e)/p(w_j)$$

$$= p(w_j|x_e) \sum_{i=1,m_s} p(x_s|d_{si},w_j) \cdot p(d_{si}|w_j)/p(x_s) \qquad (8)$$

$j = 1,2,\ldots,M$, where there are assumed to be M information classes and m_s spectral classes.

By rewriting (7) in the form

$$F_j(X) = [p(w_j)]^{-1} \cdot [p(x_s|w_j)p(w_j)/p(x_s)] \, p(w_j|x_e)$$

$$= p(x_s|w_j) \, p(w_j|x_e)/p(x_s), \qquad (9)$$

it may be seen that this classification strategy is equivalent to that described in [2]. That is, the form of the discriminant functions is essentially the product of a class-conditional probability times the probability of observing the class at the elevation at which the observation was made. The more detailed expression, equation (8), shows how the spectral classes are properly treated if, as is often the case, unsupervised analysis is used.

Notice that for this example, the assumption that the data sources are independent is likely to be reasonably well satisfied. That is, the spectral response of a given forest species may reasonably be assumed to be independent of elevation. To the extent this is not the case, the model will fail to take advantage of discriminatory information available from the dependencies.

VII. PRELIMINARY EXPERIMENT

In the near future, we will be able to explore the application of this approach to data sets containing, at minimum, a geometrically registered composite of Landsat MSS data, aircraft multispectral scanner data, side-looking radar data, topographic data (elevation and derived slope) and digitized land use maps. At this writing, however, the assembly of these data sets had not been completed. Therefore, it was decided to pursue the following experiment as a demonstration of the concepts.

A subscene (82 x 100 pixels) of a Landsat MSS image over an agricultural region of New South Wales, Australia, was analyzed using all four spectral bands. The subscene is shown in Figure 1. For the purposes of this initial experiment, it was decided to define the information classes based on spectral characteristics of the scene rather than actual ground cover; our goal was to assess the ability of the method to capture and utilize information in the data rather than to achieve an "accurate" classification per se. This will become clearer as the methods are described.

To establish a baseline result, a supervised spectral analysis was performed. Spectrally distinct regions in the image were located by displaying the subscene on a color image display system and manually selecting regions of notably different color. A Gaussian maximum likelihood classification of the entire subscene was performed based on the mean vectors and covariance matrices for these regions. By applying a light threshold to the classification discriminant values, additional regions were determined which were spectrally distinct from those already selected. The new regions were added to the old and the classification repeated. This process was iterated until virtually the entire subscene was accounted for by the accumulated spectral classes, which were eleven in number. The eleven spectral classes were defined to be the information classes, and a classification map of the area then constituted the reference against which all subsequent trial classifications would be compared.

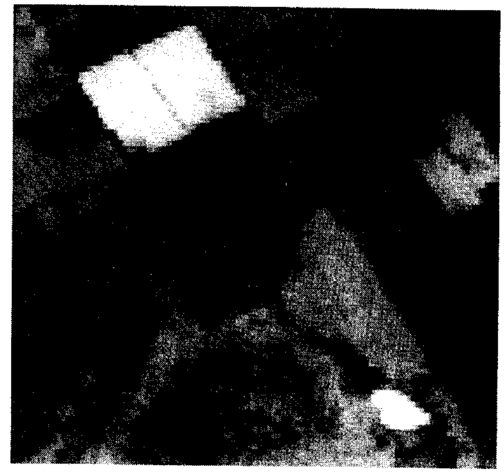

Figure 1. Subscene used for preliminary experiment.

The method for multisource data analysis set out in section IV was then applied to the same area, treating the visible bands 1 and 2 (0.5-0.6 and 0.6-0.7 micrometers) and the infrared bands 3 and 4 (0.7-0.8 and 0.8-1.1 micrometers) as two separate data sources. Table 1 shows the statistical correlations among the four bands for the subscene. Notice the relatively low correlations between pairs of bands from different spectral regions as compared to the correlations between pairs of bands from the same region. Thus we assume for this exercise that the two sources are "relatively" independent.

The analysis proceeded as follows:

1. For each data "source" (spectral region), the subscene was clustered independently to derive a set of data classes appropriate to that source. Bands 1 and 2 yielded 12 data classes; bands 3 and 4 yielded 15 data classes. Mean vectors and covariance matrices were computed for each set of data classes and

Table 1. Statistical correlations between spectral bands for the test subscene.

	Band 2 (.6-.7)	Band 3 (.7-.8)	Band 4 (.8-1.1)
Band 1 (.5-.6)	.8056	.2261	-.1473
Band 2 (.6-.7)		.3703	-.0840
Band 3 (.7-.8)			.8672

the subscene was classified, independently, based on each set.

2. In order to apply equation (4), the source-specific probabilities were cast in the form

$$p(w_j|x_s) = [p(x_s)]^{-1} \cdot \sum_{k=1,m_s} p(x_s|d_k,w_j)p(d_k,w_j) \quad (10)$$

where m_s is the number of data classes for source s (s = 1,2) and $p(x_s)$ is computed by

$$p(x_s) = \sum_{j=1,M} \sum_{k=1,m_s} p(x_s|d_k,w_j)p(d_k,w_j) \quad (11)$$

Here M is the number of information classes. The joint probabilities $p(d_k,w_j)$ were tabulated by comparing the classifications from the individual sources to the reference map. To reduce considerably the computation and memory requirements, the class-conditional probabilities were computed independently of information class; i.e., we set

$$p(x_s|d_k,w_j) = p(x_s|d_k) \quad \text{for all } w_j. \quad (12)$$

This is true if the distribution of the data within a data class is the same regardless of information class. This condition is unlikely to hold exactly, but the approximation seems to be essential to the feasibility of the computations. This is discussed further below.

3. The subscene was then classified using the global membership function defined by equation (4).

Table 2 shows the results of this composite classification as well as the results obtained from the individual data sources. Tabulated is the percent agreement with the reference map. The overall classification accuracy of the composite is substantially better than that obtained from either single spectral region. Apparently spectral class "9" in the reference classification was not isolated by the clustering algorithm applied to either of the individual spectral regions. This represents a shortcoming in the unsupervised classification method used to analyze the individual data sources rather than a problem inherent in the multisource classification approach.

VIII. DISCUSSION

The fact that the composite classification result is better overall than either of the two individual source results demonstrates only that the proposed approach for merging information from multiple sources can be successful. The difference between the composite result and the reference classification (18.5 percent) represents the degree to which the total analysis procedure used here failed to capture discriminatory information apparently contained in the four-band multispectral data. This failure may be attributed in part to each of several factors. For one thing, the reference classification was supervised while the visible and infrared classifications were unsupervised. Also, the analysis procedure based on the global membership function given by equation (4) fails to account for dependencies between the two sources (this is related to the independence assumption made in deriving this global discriminant function from the posterior probabilities). Finally, there is the approximation, equation (12), made to reduce the computation and memory requirements. Additional studies are required to assess better the impact each of these factors will have on practical application of the method.

There are some significant benefits which accrue from using the product form of the global membership function, equation (4), benefits arising principally from the decoupling of the sources in the analysis process. Most of these were mentioned earlier, in the discussion of the independence assumption. The computational complexity of the analysis process is likely to be lower than would be required if all variables had to be utilized simultaneously. This in turn means that the total amount of computer time required is likely to be less, as will the amount of

Table 2. Classification results for two data sources and the composite.

	Percent Agreement with Reference for Class											
	1	2	3	4	5	6	7	8	9	10	11	OA
visible	96.0	86.4	78.1	98.4	9.3	79.2	9.0	0	0	88.5	2.2	57.6
infrared	88.9	99.7	89.9	89.1	74.8	68.3	80.9	0	0	75.7	0	71.7
composite	95.0	99.8	94.7	92.9	76.7	92.1	85.4	45.2	0	87.4	39.7	81.5
# pixels	642	640	1221	1271	1362	518	514	598	191	863	277	8092

training data required, both of which will lower the cost of the analysis. Perhaps most importantly, each of the data sources can be dealt with on its own terms, using analysis methods only as complex as necessary for that particular source. The analyst is given the ability to reprocess selectively individual data sources without repeating the analysis of the entire ensemble of sources. Likewise, the relationships among the data classes and the information classes may be altered and the composite classification recomputed without repeating the analyses of the individual data sources.

IX. CONCLUSIONS

A general approach has been formulated to accomplish merging of information from diverse data types in geographic information systems. Key aspects of this method include:

- The definition of data classes which correspond in a natural way to each of the logically independent data sources;
- The relating of the data classes to the information classes through a set of source-specific membership functions;
- Merging of information from the individual data sources through a set of global membership functions upon which the actual classification decisions are based.

It has been shown that at least one previously successful method for handling multisource data is readily described in terms of the proposed product form of the global membership function. Preliminary experiments have demonstrated the ability of the proposed approach to merge information from separate sources.

Many aspects of the analysis of multisource geographic data remain to be addressed. Our initial method has left aside the matter of the relative quality of the respective sources; and we have made no attempt to deal here with spatial information or the different aspects of point, line and area features. These are all matters which will eventually require attention in the development of a comprehensive system for geographic information processing.

REFERENCES

[1] R. M. Hoffer and staff, "Computer-aided analysis of Skylab multispectral scanner data in mountainous terrain for land use, forestry, water resources and geological applications," LARS Information Note 121275, Laboratory for Applications of Remote Sensing, Purdue University, W. Lafayette, IN 47907, 1975.

[2] A. H. Strahler and N. A. Bryant, "Improving forest cover classification accuracy from Landsat by incorporating topographic information," Proc. Twelfth Internat. Symp. Remote Sensing of Environment, Environmental Research Inst. of Michigan, April 1978, pp. 927-942.

[3] C. F. Hutchinson, "Techniques for combining Landsat and ancillary data for digital classification improvement," Photogrammetric Engineering and Remote Sensing, vol. 48, no. 1, pp. 123-130, 1982.

[4] J. A. Richards, D. A. Landgrebe and P. H. Swain, "A means for utilizing ancillary information in multispectral classification," Remote Sensing of Environment, vol. 12, pp. 463-477, 1982.

APPENDIX: THE GLOBAL MEMBERSHIP FUNCTION

The proposed global membership function for multi-attribute data is:

$$F_j(X) = F_j(x_1, x_2, \ldots x_n) = [p(w_j)]^{1-n} \prod_{s=1,n} p(w_j | x_s)$$

Alternatively, the logarithmic form may be used:

$$F_j'(X) = (1-n) \log p(w_j) + \sum_{s=1,n} \log p(w_j | x_s)$$

The product form for the global membership function is motivated by the following considerations. A discriminant function which is natural to adopt is the posterior probability $p(w | x_1, x_2, \ldots, x_n)$. Using Bayes' formula, this may be written

$$p(w | x_1, x_2, \ldots, x_n) = p(w, x_1, x_2, \ldots, x_n) / p(x_1, x_2, \ldots, x_n) = p(x_1, x_2, \ldots, x_n | w)\, p(w) / p(x_1, x_2, \ldots, x_n)$$

If the x_i's are independent (and class-conditionally independent), then

$$p(w | x_1, x_2, \ldots x_n) = \frac{p(x_1 | w) p(x_2 | w) \ldots p(x_n | w) p(w)}{p(x_1) p(x_2) \ldots p(x_n)} = \frac{p(w | x_1) p(w | x_2) \ldots p(w | x_n)}{[p(w)]^{n-1}}$$

which is the form given in equation (4).

AUTHOR BIOGRAPHICAL DATA

Philip H. Swain is a Professor of Electrical Engineering at Purdue University, West Lafayette, Indiana. Since 1966 he has been affiliated with Purdue's Laboratory for Applications of Remote Sensing (LARS) where from 1972 to 1984 he held the post of Program Leader for Data Processing and Analysis Research. During the academic year 1984-85 he was Honorary Visiting Fellow in the School of Electrical Engineering and Computer Science and the Centre for Remote Sensing, University of New South Wales, Sydney, Australia.

Dr. Swain's research interests include image processing, pattern recognition, remote sensing and geographic information processing, artificial intelligence and advanced computer architectures for multidimensional image processing. He is widely published in these fields and is coauthor and coeditor of the text Remote Sensing: The Quantitative Approach, published in 1978 by McGraw-Hill.

John A. Richards received the degrees of B.E. and Ph.D. in 1968 and 1972, respectively, both in Electrical Engineering from the University of New South Wales.

From 1972 to 1977 he was with the Department of Electrical Engineering of the James Cook University of North Queensland, Townsville, Australia. In 1977 he joined the School of Electrical Engineering and Computer Science at the University of New South Wales, where he is currently Associate Professor. He was appointed Director of the newly established multidisciplinary Centre for Remote Sensing at the University in 1981.

Dr. Richards is a senior member of the IEEE and a fellow of the Institution of Radio and Electronics Engineers, Australia.

Tong Lee received his B.E. degree from the University of New South Wales, Sydney, Australia, in 1982, in the School of Electrical Engineering and Computer Science. He was awarded the University Medal on graduation.

Mr. Lee is presently undertaking research towards a Ph.D. in Electrical Engineering in the field of image processing and pattern recognition. His research interests also include image restoration and computer graphics.