Multispectral Land Sensing: Where From, Where To?

David Landgrebe, *Life Fellow, IEEE*
School of Electrical & Computer Engineering
Purdue University, West Lafayette, IN USA
landgreb@ecn.purdue.edu

Multispectral Land Sensing: Where From, Where To?

David Landgrebe
Professor Emeritus of Electrical and Computer Engineering
Purdue University, West Lafayette Indiana 47907-1285
landgreb@ecn.purdue.edu

Abstract. This paper begins with some brief historical comments to set the stage for a discussion of the long term potential for land remote sensing technology. This is followed by comments about what is needed to accelerate the achievement of this potential. The discussion is concluded with what concomitant development is needed with regard to a hyperspectral data analysis system.

History and Background. My first involvement with land remote sensing came in the Fall of 1965, when Prof. Ralph Shay, then Head of Purdue's Department of Botany and Plant Pathology came to the Purdue's School of Electrical Engineering to present a departmental seminar. In the seminar he put forth the then somewhat radical idea that electrical engineers and agriculturalists should begin collaborating on research to create better agriculture management information based on aerospace platforms. What he had in mind was not just science, but engineering, not simply research to create new knowledge, but research to create new knowledge that would solve a societal problem.

Remote sensing methods had been around for many years, originally based on imagery from airborne photography, but the launch of Sputnik in October 1957 greatly stimulated a number of farsighted thinkers to raise their sights even higher. Ralph had recently chaired a National Research Council study on remote sensing for agriculture, and he was seeking to form a collaboration among three universities whose skills spanned the problem to develop the technology[1].

At the time of Ralph's seminar, I had recently returned to campus from a period of work for a west coast aerospace industry. They were attempting to develop products using neural networks to derive human physiological information such as the measurement of human stress, the stages of sleep, and precursors of the onset of epileptic seizures. These are very difficult problems with highly variable characteristics from individual to individual, characteristics that I was to soon learn were similar to multispectral data of the Earth's surface. In hearing Ralph's presentation, it struck me that image processing methods would not be practical for such a problem, but spectrally based methods, which were just beginning to be investigated, might well be.

These facts are relevant even now because they bring out a number of the significant aspects of the problem that are the precursors for success in addition to others that were soon to emerge in meeting the need. That is,

> There was and is a real need for *timely* information by which to better manage agricultural, forestry, and a wide variety of other land resources. This will be discussed in greater detail later.

> The problem is quite challenging in the face of the *complex and dynamic nature* of the Earth surface cover and the often subtle spectral differences in economically significant cover types. This also will be elaborated upon below.

> The problem is an *inherently interdisciplinary* one, requiring well-coordinated, fundamental, and practical knowledge, and thus coordinated participation from an appropriate subset of engineering and Earth science disciplines.

What was envisioned for the collaboration was to have people involved who were knowledgeable in all of the engineering and Earth science disciplines involved. (See Fig. 1.) People knowledgeable in the engineering subdisciplines of optics, solid state devices, and signal processing would have as their primary focus the creation of the needed sensor, spectral measurement, and information extraction systems, while the relevant Earth scientists would create the necessary understanding of the scene variables. Persons more closely involved with the operational use of the information the system could produce would then help to mold the system into practicality that meets a societal need.

[1] The three universities involved were the University of California/Berkeley, the University of Michigan/Ann Arbor, and Purdue University/West Lafayette. Additional details of this collaboration and this period are contained in Landgrebe, David, "The Evolution of Landsat Data Analysis," (Invited), Photogrammetric Engineering and Remote Sensing, Vol. LXIII, No. 7, July 1997, pp. 859-867. This is a special issue commemorating the 25th anniversary of the launch of Landsat 1, July 1972. This paper is downloadable from http://dynamo.ecn.purdue.edu/~landgreb/publications.html

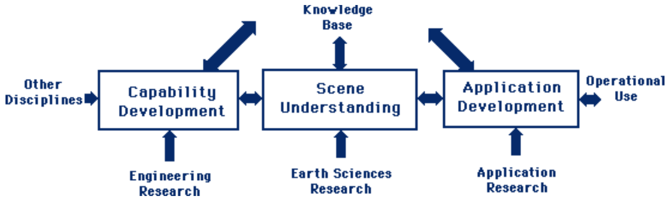

Figure 1[2]. The multifaceted nature of remote sensing technology research and development.

Indeed, these elements would need to work in close collaboration with one another. Thus for the portion of the effort at Purdue, people with each of these kinds of expertise were not only recruited to work on this problem, but they were brought together at a single office location for the entire working day so that each person's contribution could be made as relevant to the overall objective as possible and the both formal and informal interchange between them could be as thorough as possible. This was the basis of the Purdue unit originally called the Laboratory for Agricultural Remote Sensing (LARS), later changed to the Laboratory for Applications of Remote Sensing when it became apparent that the technology would have substantial value to disciplines quite beyond agriculture as Dr. Shay and the National Research Council committee had first studied.

Though some preliminary work had been underway before, work began in earnest in early 1966 with experiment design and assembling of the needed equipment and resources. The University of Michigan people had been doing surveillance technology work for the Army for some time and had constructed an airborne system based on photography for the visible and a line scanner for the middle and thermal IR. For this new work, they were able to modify this capability to provide a line scanning capability from visible though the near, middle, and thermal IR, thus making all electronic data available on analog tape. At LARS, an analog-to-digital converter was built that could digitize all spectral bands of this tape simultaneously, thus producing digital data in vector form that is the norm today. Pattern recognition methods began to be studied as the principal means for assigning these vectors to classes meaningful to the user community. First flights of the newly modified system took place in the summer of 1966, and were highly productive, producing high quality data in 12 to 18 spectral bands, data still valuable for use today at least for instructional purposes[3].

The work soon began to be taken up by other institutions and organizations. By 1968, NASA began to lay plans for a satellite to be called the Earth Resources Technology Satellite (ERTS) and later renamed Landsat. There are many stories, some true, others probably not, about how this satellite came about. One of this type was that the firm that had made imaging sensors for weather satellites suggested that it would be easy to build a similar satellite for monitoring the land by adjusting the dynamic range of the sensor from that appropriate to see (bright) clouds to that for imaging the (much darker) land surface. The telemetry system would be an analog based system suitable for producing imagery on the ground like the heritage weather systems. But the research work based on the Michigan airborne system and pattern recognition work strongly suggested the need for measurements in the infrared, and for digital data rather than just imagery. Thus, a compromise was reached to include both types of systems, the Return Beam Videcon (RBV) based imaging system and the line scan system that was called simply Multi Spectral Scanner (MSS). However, the MSS was seen as a significant disappointment to those who had been working with the multispectral airborne data. Only four spectral bands could be included in MSS, given the young state of space sensor technology at the time, whereas the precursor work had been studying the capabilities provided by many more than four. Thus MSS had to be seen as only a preliminary first step, with more capable systems to follow in later generations of sensors.

Progress on the basic technology came very rapidly after its initiation in 1966. Prior even to the launch of that first land satellite, an opportunity arose to test the state of development of the technology, its potential value to society, and to demonstrate it to a wider audience. This opportunity came in the form of a threatened national emergency to the U. S. nation's corn crop. In the latter part of the 1970 growing season, a pathogen began showing up in the southern portion of the U. S corn belt that attacked corn plants. This fungal pathogen, called Southern Corn Leaf Blight, developed from airborne spores and first showed up as brown lesions on the lower leaves of the corn canopy. The lesions would grow in size and spread upward in the canopy until ultimately, the entire plant was destroyed. The broad susceptibility of the corn crop stemmed from the fact that most varieties of corn in the U. S. were hybrids utilizing a single type of cytoplasm, and it was this cytoplasm that was the basis for the susceptibility. By the time the danger was realized, seed corn for the 1971 year had already been produced, and so it was feared that, if Southern Corn Leaf Blight spores could survive through the 1970-71 winter, it would be ready to spread rapidly during the 1971 growing season and devastate corn production.

The US Department of Agriculture mobilized several efforts over the 1970-71 winter to track the development of the possible epidemic, one of which was to try the new and untested technology called remote sensing. Since there was not yet any satellite, the test had to be done with airborne data, and since there was then only the University of Michigan multispectral scanner system mounted in a low altitude C-47 aircraft, it could not cover the whole US Corn Belt. Thus a

2 Adapted from the introduction to a special issue of the *Proceedings of the IEEE* entitled "Perceiving the Earth's Resources from Space," Vol. 73, No. 6., June 1985.

3 See for example a data set contained on the CD accompanying David Landgrebe, "Signal Theory Methods in Multispectral Remote Sensing," John Wiley and Sons, Inc, 507 pages, 2003.

high altitude photographic reconnaissance system was also made available to the collaborative team in addition to the Michigan system.

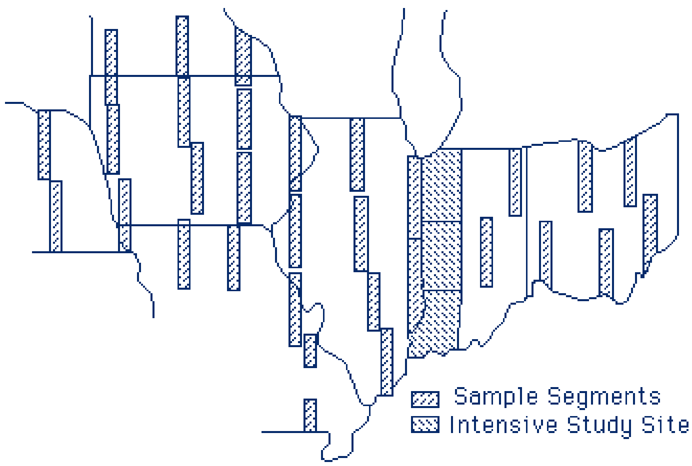

Figure 2. 1971 Corn Blight Watch Experiment sample flightlines.

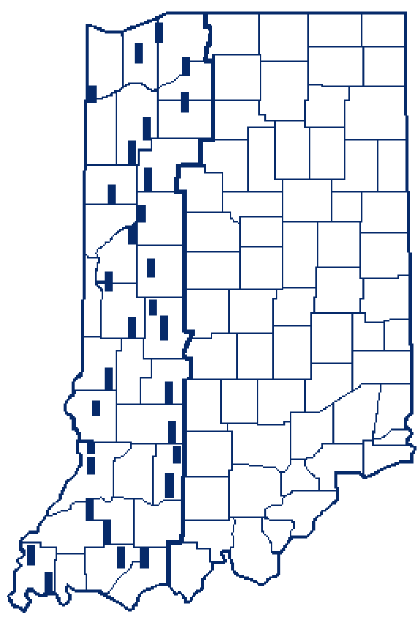

Figure 3. The sample segments of the CBWE Intensive Study Site.

Fig. 2 shows the sample segments to be flown by the photographic system, and Fig. 3 shows those to be covered by the multispectral scanner. In this way, in addition to a test of the overall technology, a comparison could be made of the relative effectiveness of the two systems, the image based one and the spectrally based one. Each sample segment was to be overflown every two weeks throughout the growing season, the data analyzed and the results reported to USDA before the next biweekly flight over that segment. As it turned out, because the aircraft assignments on any given day could be varied based on cloud conditions, no segment passes were missed due to cloud conditions during their two week period throughout the entire summer. This is a significant indication of how the cloud cover limitations could be overcome in future operational systems. It was also found that no special adjustment for atmospheric effects on the multispectral data was necessary nor was any used.

Based on a post experiment critical evaluation by skeptical USDA officials, the experiment was deemed a success. It proved possible to not only discriminate between corn and all other ground cover throughout the season, over a large land area and the various soil types, planting dates, varieties, and other agricultural variables present, but within the class corn, it proved possible to discriminate between several degrees of blight infestation, frequently even before the infestation was apparent from imagery created from the data.

The long term potential for multispectral sensing. The core potential for multispectral technology was thus well demonstrated by 1971, and indeed, it had already been developed to a practical level by that time, assuming the availability of suitable data. It is noted that data for the multispectral portion of the Corn Blight Watch Experiment included 17 spectral bands[4] from the visible blue to the thermal infrared. Indeed, the inclusion of the 9-11 μm band in the thermal region was the most selected band by the feature selection algorithm in use at that time. Its importance is possibly explained by the fact that when a plant goes into stress, an early indication is that the stress frequently results in the reduction of its canopy evapo-transpiration function resulting in a rise in the radiant thermal power emitted from the canopy.

What of the ultimate long term potential for the technology? First, a pre-assessment. In 1966 when the U.S. space effort was still in its beginning phases, NASA requested that the National Research Council undertake a study "on the probable future usefulness of satellites in practical Earth-oriented applications." A steering committee composed of top-level aerospace executives, consultants, and academicians began the work by constructing a list of areas where this use might occur. Then under the guidance of the steering committee, a major study was conducted during the summer of 1967 at the NRC's Woods Hole, MA study center. The study was carried out by a group of 13 panels, each consisting of a number of qualified persons in the fields identified by the steering committee, as follows.

[4] For details, see David Landgrebe, "Signal Theory Methods in Multispectral Remote Sensing," John Wiley and Sons, Inc, 2003, pp 80-81.

1. Forestry-Agriculture-Geography
2. Geology
3. Hydrology
4. Meteorology
5. Oceanography
6. Broadcasting
7. Point-To-Point Communication
8. Points-To-Point Communication
9. Navigation and Traffic Control
10. Sensors and Data Systems
11. Geodesy and Cartography
12. Economic Analysis
13. Systems for Remote-Sensing Information and Distribution

The Summary Report, entitled, "Useful Applications of Earth-Oriented Satellites" done by the study steering committee, was supported by individual reports generated by each of the 13 panels. Each attempted to make projections as to how space capabilities might be made to impact their area of society in the future. Now some 38 years later, a reasonable question might be, "how did they do?" Or, from another point of view taking the steering committee's foresight as 20/20, how well has the space-oriented community done in bringing about the developments they projected?

In examining the study areas, it is perhaps appropriate to note that the disciplines represented by the last four panels, Sensors and Data Systems, Geodesy and Cartography, Economic Analysis, and Systems for Remote-Sensing Information and Distribution, are best regarded as supporting areas to the other disciplines, having an impact on all of the other discipline areas. It would be difficult, therefore to assess the impact of developments in these areas separately. The other nine have been listed in groups as remote sensing of the land, atmosphere, and ocean, a communication group, and then navigation. They are also listed in approximate descending order according to their expected ultimate impact on our society, at least as seen by the participants at the Woods Hole meeting. That is, the participants appeared to believe that land remote sensing would ultimately have the greatest economic impact, since it is on the land where the people live and the major amount of economic activity takes place. Next came the atmosphere and ocean, and then the communication group; and many at the meeting in 1967 had a difficult time perceiving how satellites could impact navigation and traffic control. Comments will next be given on the present day impact of each area, starting with the one originally projected to have the least potential and moving up to the most.

Navigation and Traffic Control. One has only to think of the Global Positioning System and the fact that, for a small amount of money, anyone can buy a small device for one's car, boat, or to carry by hand that can determine with great precision one's location on the surface of the Earth. Many additional applications, precision agriculture and navigation systems for autos are two examples but there are many more that make specific use of this capability. It is indeed in common everyday use by the general public, and the steering committee should receive an A grade for anticipating this application. Or rather, the space related community should receive an A grade for its development. This is in large part thanks to the U. S. Department of Defense building the needed space infrastructure based on a fleet of more than 20 satellites.

Points-to-point Communication. Though perhaps of less utility to the general public than the others, here again, the anticipated developments have come to pass. This is certainly the case for the military, and for such other pursuits as search and rescue operations, and the tracking of animals in the wild. A widely advertised commercial product installed in personal autos makes use of such a capability.

Point-to-Point Communication. Here also, an effective technological development has taken place. All kinds of data for government and private commercial organizations and for individual users have been transmitted this way for some years.

Broadcasting. Here again, the steering committee's vision has been developed effectively. Anyone can, for a small amount of money, acquire a dish antenna and receive 150 or more channels of digital television, for example. Many channels of satellite radio are also available. This time the development was done primarily by the commercial sector, because there was already a market present.

Meteorology and *Oceanography*. These two are also clear. Indeed, the applications in these two areas were already under way at the time of the study. Anyone can turn on a television set or access an appropriate website and see images of the current weather at will. There exists a whole host of private value-added companies whose function it is to make specialized weather projections for specific individual segments of society. Oceanographic information is less widely available, but it might be fair to say that the demand or need for such is more narrowly focused. This development was undertaken primarily by the federal government; a federal agency, NOAA, was specifically placed in charge of these two areas.

Land Remote Sensing. This leaves the first three panels, the three associated with the resources and applications of the land. It would seem fair to say that, while the steering committee and the members of the 13 panels (I was the chair of the Sensors and Data Systems Panel) would be very satisfied with the developments in the other areas, they would not feel that the developments in the land remote sensing areas have yet come close to their full potential. Though there is some use in the scientific community, there are not many federal, state, or local agencies or individuals that rely on this technology as an information source on a continuing basis. Unlike the Navigation and Traffic Control and the Meteorology and Oceanographic areas, it was decided as early as the Carter administration that land area applications should be pursued by the private sector rather than the government. The development has not been vigorously pursued, because there was not a pre-existing market for use of such data. Such a market will need to be developed as the potential for the area is realized.

The Landsat system of satellites may be a useful metric in this regard. There have really been only two generations of Landsats built since it was conceived back in the 1960's. Landsat 7, a second-generation system launched April 15, 1999, contains seven of the same spectral bands at the same 30-meter spatial resolution and coverage that were committed to in 1975[5]. The limitation to only seven spectral bands substantially limits the information that can be obtained from the data, and is arguably significantly less than that obtainable from the 17 bands of the 1971 Corn Blight Watch Experiment.

Unlike the weather, communication, and navigation areas, there have been generally only one or at most two Landsats in orbit at a time over the years since the first launch in 1972. Only one in orbit at a time means that, given cloud cover conditions, one could go for months without being able to obtain data from any given site. Further, for the general user, there have been substantial delays between when the data over a given site are collected and when it can be in the user's hands. And, given the small existing user community and unlike the other areas, land data are rather expensive.

Indeed, it would not seem prudent for any agency to establish the infrastructure to rely on Landsat or any other such single satellite to meet a continuing need for information, since any one satellite can fail at any time and historically, there has been a very long delay in replacing a failed one. All of these facts are both indications that the field is underdeveloped and are also reasons why that is so.

Accelerating the Development. For purposes of stimulating discussion, I hypothesize the following as a minimum. To achieve the potential for the field, the need is for a fleet of perhaps 20 identical satellites in orbit at any given time. They would need to have sensors with a spatial resolution that is appropriate for a broad class of uses, and spectral bands that cover the optical range from the blue through the thermal infrared, divided into at least 20 and perhaps many more bands, and with a S/N adequate to justify at least 10 bit (1024 levels) data.

This number of satellites would allow for coverage of any given location every day, so that the limitations that cloud cover impose would provide users with actual coverage within at most a few days of the need in most areas and at most key times of the year. It would also allow for spares on orbit to increase reliability and inspire greater confidence.

Today, given the single satellite coverage, use of land satellite data are entirely retrospective in nature. That is, when a potential user has a problem that potentially requires satellite data, a search of existing archives must be undertaken to see if data of the location desired and relatively recent acquisition already exists. Nearly always, the user's needs must be compromised by what data are in the archives. Not many users can afford to have data gathered to order and even if so, cloud conditions may prevent acquiring the data in anything approaching an acceptable period. The user is thus at the mercy of the satellite system rather than the other way around, and this limitation eliminates large numbers of users and uses of the technology.

The need for immediacy will be illustrated with two examples from among the many that exist.

> A few years ago in central Indiana, a killing frost occurred on June 21, the first official day of summer. Thousands of acres of corn were lost or severely damaged, but how much and where? The Governor's Office was faced with whether to declare a disaster and if so how many counties were affected? The Governor quite naturally turned to the Purdue School of Agriculture to provide quantitative information by the next Friday, just a few days hence. It was Tuesday and the Dean of Agriculture turned to LARS and its remote sensing expertise. An inquiry regarding Landsat and SPOT revealed that the next Landsat pass was 10 days off and SPOT was 2 days off. As it turned out, it was cloudy on those two particular days, though space observations might have been possible on one or more of the other days. Instead, LARS personnel had to fly the entire affected part of the State on Friday in a small airplane, marking affected areas on existing topographic maps and old Landsat images, a considerably more costly, subjective, and labor intensive process. The resulting information was turned over to the Agricultural State Statistician who provided the needed information to the Dean. It showed that 32 of the 92 Indiana counties had been severely affected by the frost.

> A recent agronomy PhD student at Purdue studied whether hyperspectral data could be used to measure the magnitude and extent of hail damage in cornfields. In his PhD thesis, the student documented the following. The value of the corn crop in the U.S. is about $36 billion each year, and each year about 10% of that crop is lost to hail damage. Over $16 billion is paid in claims by insurance companies on all crops each year. These facts document that a significant insurance industry segment exists with regard to hail damage to crops. When there is such damage and an insurance claim is made, the insurance

[5] The seven bands of Landsat 4, 5, 6, and 7 were defined at a meeting in May 1975, more than 25 years ago based on the capabilities of spaceborne sensor technology at that time. An eighth panchromatic band with 15-meter spatial resolution was added for Landsat 7, and the spatial resolution of the thermal band was increased over that of previous Landsats.

adjuster observing the damage standing at the side of the field is likely to tend to minimize it, while the farm operator is likely to maximize it. A need exists to determine the extent of the loss quantitatively and objectively, and it must be done soon after the loss occurs. The student in his thesis was able to show that with hyperspectral data, this can be done, but to be adopted, a timely and reliable source of such data would be required.

There must be many tens of examples of this sort from throughout our economy that could be listed. If the data were available and the use were being made of it, there would be such substantial use of the technology that the cost of data could be minimal, just as it is in the case of weather, communication, and navigation data, thus removing another significant current barrier to achieving the potential of the technology.

But What About Hyperspectral Data. This brings us to another missing element beyond data availability, information extraction methods. This may be the most studied part of the system, both by signal processing engineers and Earth scientists. Many questions about information extraction have been studied and reported in the literature. Are maximum likelihood methods best, what about fuzzy set methods, genetic schemes, parametric methods, or nonparametric methods as the basis for discrimination between classes? Taken as a whole, the results seem to suggest that the discrimination algorithm is not the "tall pole in the tent." Any of several of these can be made to work satisfactorily in the right hands. The key problem seems to lie elsewhere rather than simply picking the right algorithm.

There are, of course, many ways to approach this problem. One is the following. Basic engineering fundamentals suggest that to obtain an optimal solution to any engineering design problem one must first properly and precisely build a quantitative model of the physical situation one is dealing with. Then one must solve that mathematical model to obtain the needed design solution. The key word in that statement seems to be *precisely*. It follows from the computer user's mantra, "Garbage in, Garbage out" that if one does not model the data set and the classes in it in feature space carefully enough, then no algorithm is going to give good results. In the remote sensing context, this means that one must determine the best mathematical description of the classes one desires, and the list of classes must be exhaustive so that the entire data set is well modeled. This quickly reduces to using whatever ancillary information one has to label an adequate number of samples, preferably in the data to be analyzed itself, to define the classes of interest. In this way one normalizes out many of the observation variables that are non-diagnostic, things like atmospheric effects, angle effects, etc, thus greatly simplifying the needed work by not needing to "correct" for them.

This matter of precision of modeling becomes even more important as the number of bands goes up. Hyperspectral data has inherently very great potential for discrimination between classes. If one has10 bit data in 100 bands, that means that in the vector space implied, there are $(1024)^{100} \approx 10^{300}$ discrete locations. That is a very large number. Even for a data set of a million (10^6) pixels thrown into such a feature space, the probability that any two will land in the same discrete location in the space is vanishingly small. One could thus conclude that, since there is no overlap of pixels in such a space for that data set, in theory, everything is separable from everything. That is the good news. But the bad news is that in a space with such a large volume, i.e., so many discrete locations, one would need to locate decision boundaries very precisely among that large number of discrete locations to realize that advantage of the hyperspace, thus the need for precision in the quantitative model of each class. Increasing numbers of bands means increasing discrimination potential, but very rapidly increasing need for class definition precision.

One is, of course, counting on the physical parameters of the various desired spectral classes to cause the pixels to tend to group themselves to some extent in various parts of the space, and again, given the great volume of such a vector space, a large part of the space may be expected to be empty for any given data set. Taken together, these two points mean that the essential structure of the classes will exist in a lower dimensional subspace. This suggests the value of so called feature extraction algorithms, algorithms that can determine in which subspace this essential diagnostic structure for the classes at hand exists. Finding this lower dimensional structure significantly reduces the sensitivity for precision of the class models, but here too, doing so effectively depends very much on how well (precisely) the desired classes are specified.

All of this suggests that very many user needs for information can be met if the spectral data, of course acquired at the time and place required by the user need, has adequate dimensionality, and that in analyzing the data, a premium is placed on establishing mathematical models of the desired classes with as much precision as possible. These facts tend to be fundamental and are the case regardless of the specific algorithms or methods of data analysis.

What if the necessary fleet of satellites does materialize, such that data can be collected as the user requires, and a rapid delivery system is implemented so that a user need can be serviced in a reasonable manner? The cost of the data would still undoubtedly be high so long as the user community is small. On the other hand, if it is to grow to be large, what would be required is an analysis procedure that any of a large and diverse user community with only layman signal processing skills could use it effectively. The analogy I have used with my students is that one should not have to understand the theory of internal combustion engines to drive a car. One just wants to get in the car, turn the key and drive anywhere the car is capable of taking you.

Following this analogy for hyperspectral data analysis, what is needed is an analysis process that is robust in the sense that it would work effectively for data of a wide variety of scenes and conditions, and can be used effectively by users rather than only by producers of the technology. The algorithms do not need to be simple, but they must be simple to apply and robust against the variety of user problems. This seems to be the challenge before us at this time if the goal is to take the technology to its true potential. We will have to count on the commercial sector to eventually provide the needed data source, since the government has not as it has for atmospheric and oceanography sensing and for the navigation and traffic control areas, but there must exist analysis procedures that work for the user when used by the user, whether that user is the final consumer of the information or a value-added producer supplying the information to that final consumer.

Summary. A great deal of progress has been made in the land remote sensing field since the launch of the first land oriented satellites. It has become possible to build spaceborne sensor systems with large numbers of spectral bands, reasonable spatial resolution, and acceptable signal-to-noise ratios; a large variety of data analysis algorithms have been investigated and tested generally on single data sets; and a wide variety of scene types have been explored. Still, the field has not reached a level of utility close to its potential or near to that of other related space based technologies. This results in data being much less available and more expensive than it should be. Even more significantly, data cannot be collected on demand by a user; instead, the user is at the mercy of cloud conditions and the limited number of vehicles in operation. The need is for a fleet of identical satellites, so that a user can obtain data over a site of interest on his/her demand and at reasonable cost. This need must await the slow but steady progress in the private sector to see enough market potential to raise the required capital to build the fleet. More germane to this journal, there is a concomitant need for a robust and effective analysis procedure that can work effectively on hyperspectral data in the hands of a user whose primary skills may lie outside the field signal processing engineering or even of remote sensing itself.

Appendix. An Approach to Hyperspectral Data Analysis

This appendix contains an outline of the approach pursued by the author and his graduate students to generate a viable system for hyperspectral analysis for the broad class of future analysts. It is offered only as an illustrative example and certainly not the only possibility, nor is it comprehensive with regard to the whole field.

If the land remote sensing field is to reach its full potential in the coming years, there will necessarily be a large number of people who will need to begin analyzing data. These people will be staff of new and existing value-added companies, staff of many national, state, and local units of government, many different individuals in their own firms, persons of many different backgrounds and interests, most without previous hyperspectral experience. Some may have strong backgrounds in signal processing engineering, but many will not. Even those with strong signal processing backgrounds will not necessarily be familiar with all the research in hyperspectral analysis that has gone on before.

Thus there will be a need for a sound hyperspectral analysis system that is robust with regard to such a community. There are many valid approaches to this problem and many contexts for it. In order to bound the problem for purposes of this brief appendix, we will outline an approach focused solely upon the analysis of a single set of hyperspectral data (as compared, for example, to multiple multi-temporal sets or sets used in conjunction with GIS data) with the objective of producing a thematic map of ground cover classes.

The ideal analysis procedure for such a community may contain complex algorithms, but the complexity must be transparent to the user. It should be easy to apply in its default mode but can have parameters that optionally can be manipulated by more knowledgeable analysts. It should be economical to run, and should be workable on widely available, inexpensive hardware. It should assume a minimum amount of preprocessing. It should produce near optimal results in terms of accuracy and appearance when applied by users of various skill levels. Obviously this is a tall order, but it represents a useful goal against which processing research can be directed and subsequently measured.

As stated above, given adequate dimensionality distributed over the spectrum, the key limitation to good performance of analysis algorithms is the precision with which the data set and the classes desired are modeled. Usually, the process is one of extrapolating from limited knowledge of which pixels in the data belong to which classes, and the analysis scheme must apply that pre-knowledge to all pixels in the data set.

Hyperspectral data are not simple. One element of the research program that has been pursued was to study and attempt to understand hyperspectral feature spaces from a fundamental standpoint[6]. This study made clear that high dimensional spaces have characteristics that are non-intuitive. Thus it was found to be wise to avoid using normal geometrical intuition, but rather to rely on rigorous mathematical determinations about class distributions in hyperspectral feature space. With that in mind, the broad outline of a processing scheme is given in Figure A.1.

[6] Jimenez, Luis, and David Landgrebe, "Supervised Classification in High Dimensional Space: Geometrical, Statistical, and Asymptotical Properties of Multivariate Data," *IEEE Transactions on System, Man, and Cybernetics*, Volume 28 Part C Number 1, pp. 39-54, Feb. 1998.

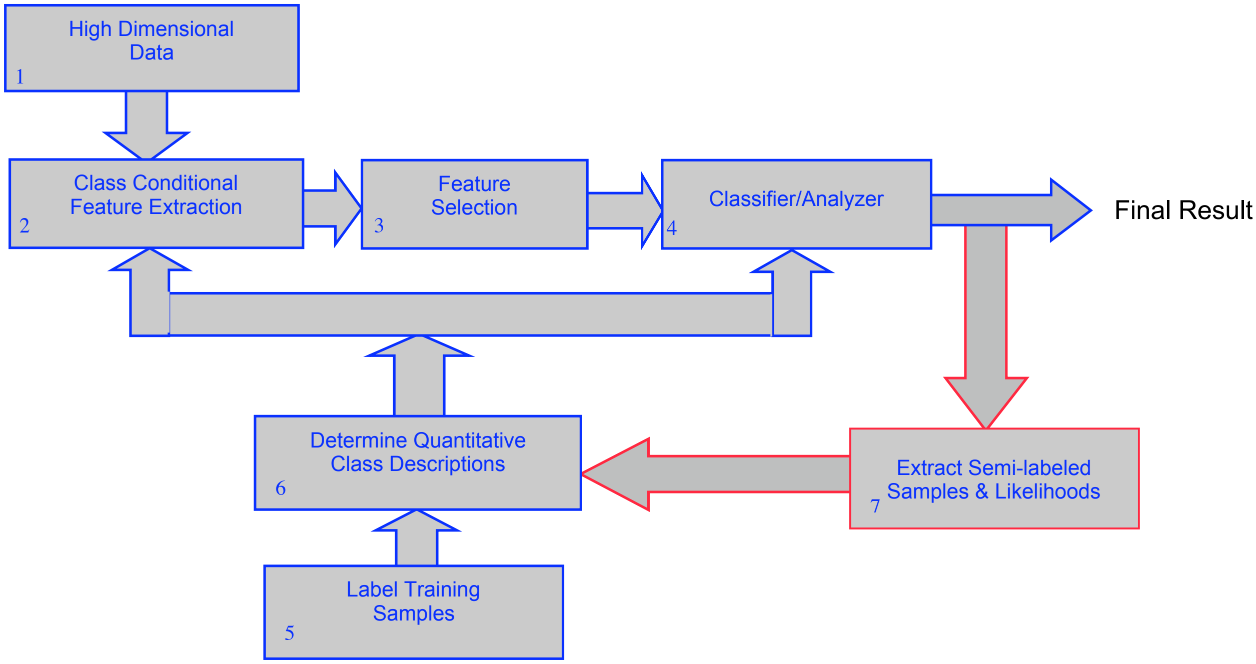

Figure A.1. The overall analysis scheme

The process begins by applying a Feature Extraction scheme (box 2 in Fig. A.1) to the data set to be analyzed to find subspaces within which the primary diagnostic characteristics lie. By doing so the portion of the entire high dimensional space that is most helpful in the discrimination process can be determined. A number of algorithms have been studied for this part of the process[7,8,9].

Feature Extraction is followed by Feature Selection (box 3), i.e. deciding on the specific subspace that is to be used, thus choosing the appropriate tradeoff between the theoretically higher discrimination power of increasing dimensionality and the rising estimation error as dimensionality increases. There follows application of the specific discrimination algorithm chosen (box 4). The most common one is maximum likelihood pixel classification, however, the possibility exists for the use of a combination of spectral and spatial information if it can be done without introducing very long classification times[10,11].

However, before the Feature Extraction and the Classifier algorithms can be applied, the model for the desired classes and the entire data set must be determined. This begins by labeling training samples (box 5), i.e. examples of each class of ground cover contained in the data set, and specifically the subset of classes of special interest to the user. This process is always unique to both the scene and the user and is the point in the process where the user's often indistinct concept of the desired classes is converted into a concrete definition. It is the beginning of the transformation from a general subjective definition to a more objective and quantitative definition. It is a part of the process that can only be adequately done by the user. To this extent, it cannot be made entirely automatic.

The means to label training samples can arise in many ways. In some cases observations from the ground may be available, for example. For another example, in the case of data of an urban area, presenting the data in image form may allow it to be done by image interpretation. In some cases, it will be possible to do so based on the shape of the spectral response of individual pixels[12].

Given the labeled samples, the next step is to create the mathematical models for the distributions of each of the classes and the data set as a whole (box 6). Since the number of training samples will necessarily be finite and usually much

7 Chulhee Lee and David A. Landgrebe, "Feature Extraction Based On Decision Boundaries," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 4, April 1993, pp. 388-400.

8 Jimenez, Luis, and David Landgrebe, "Hyperspectral Data Analysis and Feature Reduction Via Projection Pursuit, "*IEEE Transactions on Geoscience and Remote Sensing*, Vol. 37, No. 6, pp. 2653-2667, November 1999.

9 Bor-chen Kuo and David Landgrebe, "Nonparametric Weighted Feature Extraction for Classification," *IEEE Transactions on Geoscience and Remote Sensing* , Volume 42, No. 5, pp 1096-1105, May, 2004.

10 R. L. Kettig and D. A. Landgrebe, "Computer Classification of Remotely Sensed Multispectral Image Data by Extraction and Classification of Homogeneous Objects," *IEEE Transactions on Geoscience Electronics*, Volume GE-14, No. 1, pp. 19-26, January 1976.

11 D.A. Landgrebe, "The Development of a Spectral-Spatial Classifier for Earth Observational Data," *Pattern Recognition*, Vol. 12, No. 3, pp. 165-175,1980.

12 Hoffbeck, Joseph P. and David A. Landgrebe, "Classification of Remote Sensing Images having High Spectral Resolution," *Remote Sensing of Environment*, Vol. 57, No. 3, pp. 119-126, September 1996.

smaller than desirable, given the high dimensionality of hyperspectral data, the process of creating the quantitative model will be subject to significant estimation error. Thus it is important to both mitigate the estimation error as much as possible and reduce the dimensionality as far as possible while still maintaining the necessary discriminating power relative to the classes involved.

At this point also, it will not be clear what the distributions of the classes will be like. Specifically, they may not be Gaussian. It is well known that a distribution of any arbitrary shape can be modeled to arbitrary precision by using enough terms of increasing order, that is first order (mean vector), second order (covariance matrix), third order, … However, it is known that estimation error increases very rapidly with order when the number of samples available is limited. Seldom can statistics of more than second order be found usable. For non-Gaussian classes, this will not be enough.

An alternate and perhaps more practical approach is the use of multiple second order subclasses. Non-parametric schemes such as the Parzen Window Classifier are said to be "distribution free," because they do not assume any specific form of density function, but are perfectly general. Such a scheme uses a series of kernel functions to model the entire distribution, a common one being the second order (Gaussian) one. Thus, though not exactly like a Parzen classifier, one can use several second order terms to model a class and expect to achieve a good model for a complex class. This is the method that was used in the 1971 Corn Blight Watch Experiment.

However, here again, the matter of limited training becomes a factor, because the available training must be divided among the number of second order terms making up a class. More terms theoretically allows modeling to greater detail, but more terms also reduces the number of samples available with which to estimate the statistics of each and thus increases the estimation error. The question is how many terms should be used in a given case and what should their statistics be? Previously the determination as to how many terms and which training samples should be associated with each had to be done manually. If the goal boundary conditions listed above are to be achieved, a calculation procedure to determine this is needed. Such a procedure is in development[13,14,15]. Once the number of terms has been determined, there are several facets to the estimation of the statistics of each component that can be used. Some of these will be mentioned in the following.

An additional question that arises in the modeling process is how typical the training samples chosen for a given class are of that class over the entire data set. A procedure called "statistics enhancement[16]" has been created to deal with this question. It is an iterative procedure that uses the training samples and a uniform sampling of the samples of the data set not used for training to adjust the class statistics so that the composite of the class statistics better model the entire data set. In the process, because many of the non-training samples are used, this in effect increases the training set size, and estimation error is reduced. There are several other procedures that have been studied for controlling the estimation error in the face of limited training sets or otherwise improving the class modeling process[17.18,19].

One additional scheme has been investigated to reduce the negative effect of estimation error and imprecise data modeling, that of using feedback[20,21] (box 7 in Fig. A.1). In maximum likelihood classification, the likelihood of each pixel is calculated in order to assign pixels to the most likely class. Basically the idea is to choose pixels that were classified on the first iteration with high likelihood, referred to as semi-labeled samples, to augment the training set and then repeat the modeling and classification process in a second iteration. This process can be repeated for as many times as necessary in order to obtain good results. The increase in size of the training set by including semi-labeled samples reduces the estimation error and thus improves the classification process.

These elements then comprise what might become an analysis procedure of the type needed as described above. However, though the various elements of the procedure have been investigated, they have not yet been combined into an overall system that could be used successfully by a novice. And much testing of the system would be required before it could be considered to be complete. Most of the algorithms discussed here are described and illustrated in more detail in the textbook previously referred to. Many, but not all of the

---

13 M. Murat Dundar and David Landgrebe, "A Model Based Mixture Supervised Classification Approach in Hyperspectral Data Analysis," IEEE Transactions on Geoscience and Remote Sensing, Volume 40, No. 12, pp 2692-2699, December 2002.

14 M. Murat Dundar and David Landgrebe, "A Cost-effective Semi-supervised Classifier Approach with Kernels," *IEEE Transactions on Geoscience and Remote Sensing*, Volume 42, No. 1, pp 264-270, January, 2004.

15 M. Murat Dundar and David Landgrebe, "Toward an Optimal Supervised Classifier for the Analysis of Hyperspectral Data," *IEEE Transactions on Geoscience and Remote Sensing,* Volume 42, No. 1, pp 271-277, January, 2004.

16 Behzad M. Shahshahani and David A. Landgrebe, "The Effect of Unlabeled Samples in Reducing the Small Sample Size Problem and Mitigating the Hughes Phenomenon," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 32, No. 5, pp. 1087-1095, September 1994.

17 Byeungwoo Jeon and David Landgrebe, "Partially Supervised Classification Using Weighted Unsupervised Clustering," *IEEE Transactions on Geoscience and Remote Sensing*, Volume 37, No. 2, pp 1073-1079, March 1999.

18 Tadjudin, Saldju and David Landgrebe, "Covariance Estimation With Limited Training Samples," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 37, No. 4, pp. 2113-2118, July 1999.

19 Tadjudin, Saldju and David Landgrebe, "Robust Parameter Estimation for Mixture Model," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 38, No. 1, pp. 439-445, January 2000.

20 Qiong Jackson and David Landgrebe, "An Adaptive Classifier Design for High Dimensional Data Analysis with a Limited Training Data Set," *IEEE Transactions on Geoscience and Remote Sensing*. Vol. 39, No. 12, pp. 2664-2679, December 2001.

21 Qiong Jackson and David Landgrebe, "An Adaptive Method for Combined Covariance Estimation and Classification," *IEEE Transactions on Geoscience and Remote Sensing*, Volume 40, No. 5, pp 1082-1087, May 2002.

algorithms have been implemented into a software system for personal computers. This system, called MultiSpec is available for download and use by anyone without charge at http://dynamo.ecn.purdue.edu/~biehl/MultiSpec/. Additional references about this research program and downloadable copies of many of the references listed are available at http://dynamo.ecn.purdue.edu/~landgreb/publications.html.

David A. Landgrebe

Dr. Landgrebe holds the BSEE, MSEE, and PhD degrees from Purdue University. He is Professor (Emeritus) of Electrical and Computer Engineering at Purdue University. His area of specialty in research is communication science and signal processing, especially as applied to Earth observational remote sensing.

He was President of the IEEE Geoscience and Remote Sensing Society for 1986 and 1987 and a member of its Administrative Committee from 1979 to 1990. He received that Society's Outstanding Service Award in 1988. He is a co-author of the textbook, *Remote Sensing: The Quantitative Approach (1978)*, and a contributor to the *ASP Manual of Remote Sensing (1st edition, 1974)*, and the books, *Remote Sensing of Environment, (1976), Information Processing for Remote Sensing (1999)*, and *Introduction to Modern Photogrammetry* (2001). He is the author of the textbook *Signal Theory Methods in Multispectral Remote Sensing* (2003). He has been a member of the editorial board of the journal, *Remote Sensing of Environment*, since its inception in 1970.

Dr. Landgrebe is a Life Fellow of the Institute of Electrical and Electronic Engineers, a Fellow of the American Society of Photogrammetry and Remote Sensing, a Fellow of the American Association for the Advancement of Science, and the American Society for Engineering Education, as well as Eta Kappa Nu, Tau Beta Pi, and Sigma Xi honor societies. He received the NASA Exceptional Scientific Achievement Medal in 1973 for his work in the field of machine analysis methods for remotely sensed Earth observational data. In 1976, on behalf of the Purdue's Laboratory for Applications of Remote Sensing, which he directed, he accepted the William T. Pecora Award, presented by NASA and the U.S. Department of Interior. He was the 1990 individual recipient of the William T. Pecora Award for contributions to the field of remote sensing. He was the 1992 recipient of the IEEE Geoscience and Remote Sensing Society's Distinguished Achievement Award and the 2003 recipient of the IEEE Geoscience and Remote Sensing Society's Education Award. In 2003, the IEEE Geoscience and Remote Sensing Society held a workshop in his honor entitled "Advances in Techniques for Analysis of Remotely Sensed Data." Dr. Landgrebe was elected a member of the National Academy of Engineering in 2005.