1-1-1979

An Analytical Approach to the Design of Spectral Measurements in the Design of Multispectral Sensor

Daniel J. Wiersma

David A. Landgrebe

Follow this and additional works at: http://docs.lib.purdue.edu/lars_symp

Reprinted from

Symposium on

Machine Processing of

Remotely Sensed Data

June 27 - 29, 1979

The Laboratory for Applications of
Remote Sensing

Purdue University
West Lafayette
Indiana 47907 USA

IEEE Catalog No.
79CH1430-8 MPRSD

Copyright © 1979 IEEE
The Institute of Electrical and Electronics Engineers, Inc.

AN ANALYTICAL APPROACH TO THE DESIGN OF SPECTRAL MEASUREMENTS IN THE DESIGN OF MULTISPECTRAL SENSOR

DANIEL J. WIERSMA
Minneapolis Honeywell

DAVID A. LANDGREBE
Purdue University

ABSTRACT

The purpose of the research which led to this paper is to develop an analytical procedure for the design of the spectral channels for multi-spectral remote sensor systems. An optimum design based on the criterion of minimum mean-square representation error using the Karhunen-Loeve expansion was developed to represent the spectral response functions from a stratum. From the overall pattern recognition system perspective the effect of the representation accuracy on a typical performance criterion, the probability of correct classification, is investigated. Although the analytical technique was developed primarily for the purpose of sensor design it was found that the procedure has potential for making important contributions to scene under-standing. It was concluded that spectral channels which have narrow bandwidths relative to current sensor systems may be necessary to provide adequate spectral representation and improved classification performance. The optimum sensor design provides a standard against which sub-optimum operational sensors can be compared.

I. INTRODUCTION

A pattern recognition system as used in a remote sensing system for earth resources consists of three fundamental components – the scene, the sensor, and the processor (Figure 1). The scene is that portion of the earth's surface observed by the sensor. The desired information is con-tained in the spectral, spatial, and temporal variations of the electromagnetic energy emanating from the scene. The sensor collects the energy and measures its features. The processor is typically a digitally implemented classification algorithm which makes an appropriate decision based on the feature measurements provided by the output from the sensor. Various types of ancillary data are also now typically used in the decision making process.

At present the design of the processor algorithms is quite advanced and provides variety and flexibility for optimal performance given a feature set (Fukunaga, 1972; Duda and Hart, 1973). However, the design of the best set of features is a complex matter which is not well understood. In the current work we limit considerations to the design of the spectral aspects of features to make the problem more tractable, leaving other aspects to later occasions. We will attempt herein to advance the knowledge of the spectral representation of the scene and to provide an analytical basis for the design of spectral channels for an operational sensor.

II. SPECTRAL REPRESENTATION AND OPTIMUM SENSOR DESIGN

The multispectral scanner is a complex system that is designed infrequently and has very little built-in flexibility. It must serve a wide variety of users with little opportunity for specialization. As a result careful consid-eration must be given to the specification of the parameters in the sensor design. The para-meters may be grouped into five categories – spectral representation, spatial representation, signal-to-noise ratio, ancillary data or infor-mation not contained in the spectral response function itself, and the information classes desired (Landgrebe, 1978). As we have indicated, the principal category of parameters dealt with here is that of the spectral representation, since the goal is to select spectral bands; however, all of the parameters are interrelated such that specifying one of them places con-straints on the others.

The ultimate goal is to optimize the overall system performance with respect to some criterion. A typical criterion, and the one used here, is the probability of correct classification, an intricate function of the parameters.

Let us begin by considering the information bearing aspects of the spectral response function $x(\lambda)$ (Holmes and MacDonald, 1969). This response function is proportional to the electromagnetic energy received by the sensor as a function of wavelength λ (Figure 2). Many factors determine the spectral response function for a given observation. The irradiance of the sun, the conditions of the atmosphere, and the reflectance of the surface features are all known to have important effects on the response. Since a deterministic relationship between the response function and the factors affecting it would be very complex, the set of functions which are received will be modeled as a stochastic process.

The ensemble of the stochastic process (Papoulis, 1965) will be defined in terms of the stratification necessary to apply pattern recognition methods to the earth observational problem. A stratum, S, is defined as the largest contiguous area which can be classified to an acceptible level of performance with a single training of the classifier. It is noted that the sensor must be designed to operate satisfactorily over a large number of such strata, which vary greatly with time, location and application. The collection of all possible strata which a sensor may observe is denoted by S_o. Since the set S_o is quite large, it is necessary to select a smaller subset which is representative in a statistical sense in order to perform the analysis.

The random experiment for the stochastic process consists of the observation of a point in a stratum S. Each point in the stratum is mapped into a spectral response function (Figure 3). The collection of all response functions from a stratum defines an ensemble. The ensemble plus the corresponding probability measure defines the stochastic process (Papoulis, 1965). It is appropriate to assume for this process a Gaussian probability measure (Crane et al, 1972).

It is necessary to choose a mathematical model for the sensor to represent the spectral response function for each observationa. Let the sensor consist of a set of N filter functions or basis functions $\{\phi_i(\lambda)\}$ such that the output of each filter is given by (Figure 4)

$$x_i = \int_\Lambda x(\lambda)\phi_i(\lambda)\,d\lambda \tag{1}$$

The output of the sensor model is a sequence, x_1, x_2, ..., x_N, which represents the spectral response by the approximation

$$x(\lambda) \approx x_1\phi_1(\lambda) + x_2\phi_2(\lambda) + \ldots + x_N\phi_N(\lambda) = \hat{x}(\lambda)$$

$$= \sum_{i=1}^{N} x_i\phi_i(\lambda) \tag{2}$$

For a small number of simple functions, the approximation is very coarse (Figure 5). It is desired to choose a (ordered) set which will be optimum in some sense over a stratum. For the purpose of achieving an optimal set, no restrictions will be placed on the physical

realizability of $\{\phi_i(\lambda)\}$.

A key consideration is the choice of the criterion for optimality. Because the sensor must function over a varied collection of strata using any of a large collection of classifiers, a criterion was chosen which is a measure of the fidelity with which the output of the sensor represents the input. We will choose the set $\{\phi_i(\lambda)\}$ such that for a given $\{x(\lambda)\}$ the approximation $\hat{x}(\lambda)$ is as close as possible to the true spectral response function. Since the uses and therefore the specific spectral attributes needed by the various users cannot be predicted, this approach insures that all of the information in $x(\lambda)$ will still be available in and recoverable from the $\{x_i\}$ produced by the sensor. A common criterion for representation accuracy is the expected mean-square representation error given by

$$E\{\varepsilon_r\} = E\left\{\int_\Lambda [x(\lambda) - \hat{x}(\lambda)]^2\,d\lambda\right\} \tag{3}$$

However, it is desirable at this point to generalize this criterion by introducing a weight function $w(\lambda)$ on the spectral interval. As will be seen presently, the weight associated with each λ can be used to introduce a priori knowledge concerning the spectrum into the analysis. Thus equations (1) and (3) become, (Wiersma, 1979).

$$x_i = \int_\Lambda x(\lambda)\phi_i(\lambda)w(\lambda)\,d\lambda \tag{1a}$$

$$E\{\varepsilon_r\} = E\left\{\int_\Lambda [x(\lambda) - \hat{x}(\lambda)]^2 w(\lambda)\,d\lambda\right\} \tag{3a}$$

We want to choose the set of basis functions $\{\phi_i(\lambda)\}$ which is optimum with respect to the spectral representation criterion of expected mean-square error $\bar{\varepsilon}_r$. More specifically, it is desired the representation be complete in the sense that the expected mean-square error for any function in the ensemble be made arbitrarily small simply by including enough terms, that convergence of the approximation to the original response be rapid in the first few terms, and that the basis functions be orthogonal to each other.

A technique for determining the set of optimum basis functions for an ensemble which satisfies the desired properties is based on the weighted Karhunen-Loeve expansion. (Davenport and Root, 1958; Van Trees, 1968; Ash, 1967; Wiersma, 1979). The solution to the homogeneous linear integral equation

$$\gamma_i\phi_i(\lambda) = \int_\Lambda K(\lambda,\xi)\phi_i(\xi)w(\xi)\,d\xi \tag{4}$$

with the covariance function of the stochastic process, $K(\lambda,\xi)$, as kernel is a set of eigenfunctions $\{\phi_i(\lambda)\}$ with corresponding eigenvalues

γ_i. If the eigenvalues are arranged in descending order, the corresponding sequence of eigen-functions can be used to form a sequence of linear combinations of the eigenfunctions which converges to the original spectral response function with arbitrarily small expected mean-square error. Furthermore, because of the ordering of the eigenvalues, convergence in the first few terms is very rapid. This rapid convergence allows one to truncate the series expansion after a finite number of terms N with mean-square error minimized over all possible choices of N basis functions. The mean square error is given by

$$E\{\epsilon_r\} = \sum_{i=N+1}^{\infty} \gamma_i \qquad (5)$$

Since the Karhunen-Loeve expansion is a well-studied technique and satisfies the desired properties for finding the basis functions, it provides a sound analytical method for determining the optimum set of basis functions.

It became apparent during this research that it may be advantageous to incorporate certain a priori information about the spectral interval into the analysis. The generalization of the Karhunen-Loeve expansion to include a weight function provided a convenient and appropriate means to incorporate such a priori knowledge into the design process.

The optimum sensor design problem may be solved on a digital computer using empirical data taken by field measurements. Some approximations must be made in order to take into consideration some practical constraints. First the response functions are not available as continuous functions but are obtained in the field by sampling the spectrum with an instrument that uses very narrow spectral windows. Secondly, the parameters of the process are not known a priori; hence, it is necessary to estimate the mean and covariance functions using a representative sample from the ensemble. Finally, because the data will be stored and processed digitally it is necessary to quantize the amplitude of the response at each of the spectral sample points. Each of these constraints can potentially contribute to the representation error. It has been shown that with reasonable care in selecting a sufficiently high spectral sampling rate, a large enough sample from the ensemble, and a large number of quantization intervals that the contribution of these factors to the representation error is small (Wiersma, 1979). The integral equation (4) becomes the matrix equation

$$\Phi\Gamma = KW\Phi \qquad (6)$$

where Φ is the matrix of eigenvectors, Γ is the diagonal matrix of eigenvalues, K is the covariance matrix and W is the diagonal matrix of weight coefficients.

III. RELATIONSHIP BETWEEN THE SPECTRAL REPRESENTATION AND SYSTEM PERFORMANCE

The performance of the overall system is ultimately what we wish to optimize. For remote sensing problems, an often used criterion for performance is the probability of correct classification P_c. If the vector X is an observation from one of M classes C_i, $i = 1,2,...M$ with a priori probabilities P_i, the probability of correct classification, using the maximum likelihood rule is given by

$$P_c = \int_\Omega \max_i \{P_i p(X|C_i) dX \qquad (7)$$

where $p(X|C_i)$ is the conditional joint probability density function for class i. The integral in (7) is over the observation space Ω.

The analytical procedure based on the weighted Karhunen-Loeve expansion has prescribed a sensor design which minimizes the mean-square error. One would like to know how the ability to represent a process influences the classification performance. To study this relationship the graph of the probability of correct classification vs. expected mean-square error is introduced (Figure 6). We will briefly discuss some of its characteristics.

The addition of terms to the series expansion causes a decrease in the spectral representation error, but the effect of the additional terms on the overall system performance has to be determined. It can be shown that increasing the number of terms in the representation will never decrease the performance provided that the stochastic process is completely known. If after N terms the improvement in performance is small compared to the reduction in representation error, then the representation is sufficient. This is illustrated by case A of Figure 6 in which the threshold T indicates the minimum required $E[\epsilon_r]$. However, if the performance is showing significant improvement for a small decrease in the mean-square error, case B of Figure 6, more terms are necessary to complete the representation.

Since the parameters of the stochastic process must be estimated from a sample of the ensemble, the effect of the size of the sample relative to the dimensionality of the system becomes important. Hughes (1968) has shown that if the sample size is too small, the classification performance may actually be degraded by adding terms to the expansion. Thus it is necessary to maintain a large set of sample functions from which to estimate the statistics.

The choice of information classes also in-
fluences the performance of the pattern recogni-
tion system. For purposes of classifying the data
into distinct classes it is required that the
class list have the following properties simul-
taneously (Landgrebe, 1978):
    -Each class must be of interest to the
    user, i.e. of informational value.
    -The classes must be separable in terms
    of the features available.
    -The list must be exhaustive i.e. there
    must be a class to which it is logical to
    assign each pixel in the scene.
The classes may be arranged in a hierarchial tree
structure such that classes deeper in the tree
require more accurate representation to achieve
a given level of classification performance.

The area of the ground resolution element,
which is determined by the instantaneous field-
of-view (IFOV), the altitude of the sensor, the
scan rate, and the velocity of the sensor, are
examples spatial representation parameters. The
size of the objects which can be identified and
the energy available are influenced by the choice
of ground resolution element size. If a typical
object which one wishes to identify is smaller
than the ground resolution element size, then,
it is very difficult to classify that object.
Mobasseri (1978) has investigated the effect of
the area of the resolution element on classifi-
cation performance. Increasing the area often
improves the signal-to-noise ratio which in turn
improves the classification performance.

For a given remote sensing problem the signal
is the part of the received response which is
information bearing, and the noise is that part
which is non-information bearing. The influence
of the signal-to-noise ratio where the noise is
white, Gaussian and additive was demonstrated
by (Ready et al, 1971). Results show that over-
all classification performance decreased with an
increase in the noise level. A class which was
difficult to identify under low noise level
conditions suffered the most degradation when
noise was added.

IV. EXPERIMENTAL SYSTEM

An experimental software system has been set
up to implement the analytical procedure that
has been developed. The software system has
been implemented on an IBM 370 computer at the
Laboratory for Applications of Remote Sensing
(LARS) at Purdue University.

A collection of field data consisting of
spectral response functions on three dates from
Williams County, North Dakota and three dates
from Finney County, Kansas was available from
the field measurements library at LARS. More
than one thousand spectra were available from
each location and collection date. The response
functions were sampled in wavelength using
narrow windows of .02 μm.

The optimum set of basis functions is found
numerically by estimating the covariance matrix
from the sample response function. Maximum like-
lihood estimates of the mean and covariance
matrices are given by

$$\overline{X} = E\{X\} = \frac{1}{N_s} \sum_i^{N_s} x_i \qquad (8)$$

and

$$K = \frac{1}{N_s} \sum_{i=1}^{N_s} (x_i - \overline{X})(x_i - \overline{X})^T \qquad (9)$$

where N_s is the number of sample functions avail-
able and x_i is the ith sample vector. The co-
variance is the kernel in the linear integral
equation whose solutions are the optimum basis
functions or eigenvectors. A remarkably stable
and accurate method of numerically computing the
eigenvalues and eigenvectors was published by Grad
and Brebner (1968).

The eigenvectors are used to perform the
linear transformation

$$y_i = \phi_i^T (X - \overline{X}) \qquad (10)$$

on the data set. The class conditional statistics
are computed using the transformed data.

In order to compare the performances of two
systems an algorithm which estimates the prob-
ability of correct classification for an M class
problem given the class conditional multivariate
Gaussian statistics was used (Wiersma, 1979).
This algorithm, which is based on the stratified
posterior estimator, (Whitsitt and Landgrebe,
1977) was found to be accurate within one-half
of one percent.

The experimental system also included an
ability to simulate (suboptimal) practical sensors.
Although nearly any sensor characteristic could
be simulated, most of the sensors which were
simulated consisted of a small set of rectangular
basis functions, i.e.

$$\psi_k(\lambda) = \begin{cases} 1.0 & \lambda_k \leq \lambda \leq \lambda_{k+1} \\ 0.0 & \text{otherwise} \end{cases} \qquad (11)$$

where the λ_k are endpoints of the spectral
channels. The endpoints of two suboptimum sensors
which were implemented are listed in Table 1.

V. RESULTS

One of the first tasks in using the spectral
parameter design system was to select a weight
function. The uniform weight function of Figure 7
was tried first implying that no knowledge about
signal and noise regions is available a priori.
Plots of the first four eigenvectors for this
weight function are shown in Figure 8. It was
observed that the eigenvectors were dominated in
several cases by components in the bands near

Table 1. Spectral Band Locations for Two
Practical Sensor Designs.

Sensor Number 1		Sensor Number 2	
Band	Wavelength	Band	Wavelength
1	.5 to .6μm	1	.45 to .52μm
2	.6 to .7μm	2	.52 to .60μm
3	.7 to .8μm	3	.63 to .69μm
4	.8 to 1.1μm	4	.76 to .90μm
		5	1.55 to 1.75μm
		6	2.08 to 2.35μm

1.4 and 1.9 micrometers. These bands are domi-
nated by water absorption in the atmosphere, and
the eigenvectors which were sensitive to these
bands contributed very little to the classi-
fication performance. The second weight function
which assigns a very small weight in the water
absorption bands was, then, tried. (Figure 9).
The influence of the spectrum in the water
absorption bands is significantly reduced, and
a marked improvement in classification perform-
ance was observed from the first few eigenvectors.

Results for one of the six sets of data are
presented in Figures 10 through 12. The data
were collected over Williams County, North Dakota
on June 29, 1977. In Figure 10 the expected
mean-square error is plotted as a function of the
number of terms in the expansion. The rapid
convergence in the first few terms is demonstrated
by this graph. The first eight weighted eigen-
vectors are plotted in Figure 11. Examination
of the eigenvectors will provide some at least
subjective indications as to which spectral
aspects are important to system performance. It
is interesting to note that magnitude of the
first eigenvector corresponds to the average
response over the ensemble. The three classes
represented in this data set were spring wheat,
summer fallow and pasture. Since the data are
from relatively early in the season, the wheat
canopy is still quite thin and thus two of the
three classes are spectrally similar to bare
soil. This response is reflected in $|\phi_1(\lambda)|$.
In the second eigenvector there begins a tendancy
to represent the spectra in bands. As the number
of terms in the expansion is increased, the terms
that are added require higher spectral resolution
to reduce the mean-square error. In Figure 12 the
probability of correct classification is graphed
as a function of the expected mean-square error.

Although the primary purpose of this work is
to produce a design procedure for sensors, im-
portant contributions to the understanding of the
scene can be gained. Properties such as di-
mensionality, maximum possible classification
performance, spectral resolution and the accuracy
of spectral representation required to obtain a
given level of performance can be studied.

The dimensionality can be defined as the
number of terms necessary to represent the
original waveform to the desired accuracy. For

the information classes used on the six data sets
the dimensionality was about six to eight. Some
strata required fewer terms to obtain a good
representation and good performance, while others
required more terms.

The graph of classification performance as
a function of expected mean-square error proved
valuable for studying the relationship between
representation accuracy and classification per-
formance for the family of functions $\{\phi_i(\lambda)\}$. It
is possible to see which terms contributed to both
the representation and the classification per-
formance. By examining the asymptotic performance
as $E\{\varepsilon_r\}$ approached zero, estimates of the maxi-
mum classification performance were made. For
the June 29 data set the maximum value of prob-
ability of correct classification was about 0.96,
which is the value for 10 optimal features.

Each of the optimum basis functions was
ranked according to its ability to classify points
in the stratum. The ranking was based on single
feature classification performance and multiple
feature performance using divergence calculations
and the performance estimator. The rankings for
the June 29 data set are in parenthesis in
Figure 11. It was found that the ranking on
performance is not drastically different from
the ranking based on representation accuracy.
In general the first five eigenvectors tend to
have a higher ranking than eigenvectors six
through ten.

The eigenvectors were used to evaluate the
spectrum to determine effective ways of sampling.
It was observed that the first few eigenvectors
had relatively wide subintervals. Eigenvectors
later in the sequence exhibited high frequency
(i.e. high spectral resolution) variations in
magnitude. Examining several of the performance
vs. representation curves it was found that these
later terms were often significant thus indicating
a need for high resolution sampling bands in
certain portions of the spectrum. In particular
it was found that narrow bandwidths on the order
of .02 to .05 micrometers were needed in the
visible region between 0.60 and 0.70μm

The primary purpose of designing an optimal
sensor was to use it as a standard to compare
and evaluate practical sensor systems. With the
present state-of-the-art in sensor system design
functions such as those in Figure 11 are not
practical. Practical designs such as the ones
described by equation 11 and Table 1 can be
evaluated and their performances compared to the
optimal system. Based on the comparisons, one
can make a decision as to whether the candidate
design is satisfactory or some modifications
need to be made.

The comparisons between the two simulated
sensors of Table 1 and the optimum set were made
on the basis of representation accuracy (Table 2)
and classification performance (Figure 13).

Sensors 1 and 2 provide much poorer representation accuracy chiefly because they do not represent the whole spectral interval. For the information

Table 2. Comparison of Expected Mean-Square Error (in relative units) for Data Taken Over Williams County, North Dakota on June 27, 1977

Sensor 1	17320
Sensor 2	16380
First 4 optimal basis functions	26.31
First 6 optimal basis functions	11.37
First 10 optimal basis functions	5.253

classes used for this data set much of the information in the response is not required for discrimination between the classes. The improvement in representation by sensor 2 over sensor 1 is very significant. The optimal sensor consists of the first 10 eigenvectors, where 10 provides a small mean-square error yet keeps the computational problems involved with high dimensional systems to a reasonable level. As shown in Figure 13 the four band sensor 1 compares very poorly with the best four optimal ϕ_i, showing sensor 1 to be far from optimal for these classes. On the other hand the six band sensor 2 compares very favorably with the best six optimal ϕ_i, the difference in estimated P_c being certainly less than the uncertainty of P_c estimation. The slight further improvement of the first ten optimal ϕ_i is shown for comparison. However, since there is a large difference in representation accuracy, this favorable comparison may not be true for a different set of information classes.

VI. CONCLUSIONS

An analytical procedure has been developed for the design of the spectral characteristics for a multispectral remote sensor system. The procedure provides a standard for comparison of suboptimum sensors and a basis for the selection of spectral channels for operational sensor systems.

Significant contributions to the understanding of the scene were gained from this investigation. In particular knowledge about important spectral bands was extracted and efficient spectral sampling techniques were developed. Determination of scene properties such as signal dimensionality and maximum probability of correct classification was performed.

VII. REFERENCES

Ash, R. B., (1967), Information Theory, Wiley Interscience, New York.

Davenport, W. B., and W. L. Root, (1958), An Introduction to the Theory of Random Signals and Noise, McGraw-Hill, New York.

Crane, R. B., W. A. Malila, and W. Richardson, (1972), "Suitability of the normal density assumption for processing multispectral scanner data," IEEE Trans. Geosci. Elect. Vol. GE-10, pp. 158-165.

Duda, R. O., and P. E. Hart, (1973), Pattern Classification and Scene Analysis, Wiley, New York.

Fukunaga, K., (1972), Introduction to Statistical Pattern Recognition, Academic Press, New York.

Grad, J. and M. A. Brebner, (1968), "Eigenvalues and eigenvectors of a real general matrix, Algorithm 343," Association for Computing Machinery, Comm., Vol. 11, No. 12, pp. 820-826.

Harnage, J. L., and D. A. Landgrebe, (1975), "Landsat-D Thematic Mapper Technical Working Group - Final Report," NASA/Johnson Space Center Report No. JSC-90797, Houston, Texas.

Holmes, R. A., and R. B. MacDonald, (1969), "The physical basis of system design for remote sensing in agriculture," Proceedings IEEE, Vol. 57, pp. 629-639.

Hoffer, R. M., (1978), "Biological and Physical Considerations in Applying Computer-Aided Analysis Techniques to Remote Sensor Data," in Remote Sensing: The Quantitative Approach, P. H. Swain and S. M. Davis, Eds., McGraw-Hill, New York.

Hughes, G. F., (1968), "On the mean accuracy of statistical pattern recognizers," IEEE Trans. Inform. Theory, Vol. IT-14, No. 1, pp. 55-63.

Landgrebe, D. A., (1978), "Useful information from multispectral image data: another look," in Remote Sensing: the Quantitative Approach, P. H. Swain and S. M. Davis, Eds., McGraw-Hill, New York.

Mobasseri, B. G., C. D. McGillem, P. E. Anuta, (1978), "A parametric multiclass Bayes error estimator for the multispectral scanner spatial model performance evaluation," LARS Technical Report 061578, Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Ind.

Papoulis, A., (1965), Probability, Random Variables and Stochastic Processes, McGraw-Hill

Ready, P. J., P. A. Wintz, S. J. Whitsitt, and D. A. Landgrebe, (1971), "Effects of compression and random noise on multispectral data," Proceedings of the 7th International Symposium on Remote Sensing of the Environment, Ann Arbor, Michigan.

Whitsitt, S. J., and D. A. Landgrebe, (1977), "Error estimation and separability measures in feature selection for multiclass pattern recognition," LARS Information Note 082377, LARS, Purdue, West Lafayette, Ind.

Wiersma, D. J., and D. A. Landgrebe, (1979),
"The Analytical Design of Spectral Measurements
for Multispectral Remote Sensor Systems," LARS
Technical Report #122678, Purdue University, West
Lafayette, Ind.

VanTrees, H. L., (1968), Detection, Estimation,
and Modulation Theory - Part I, Wiley, New York.
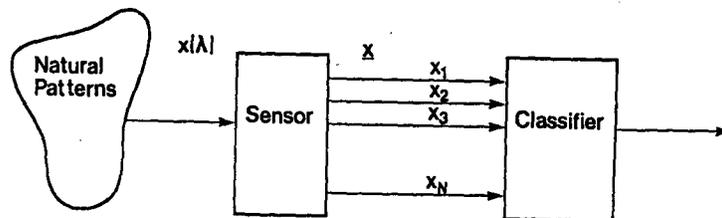
Figure 1. Pattern Recognition System
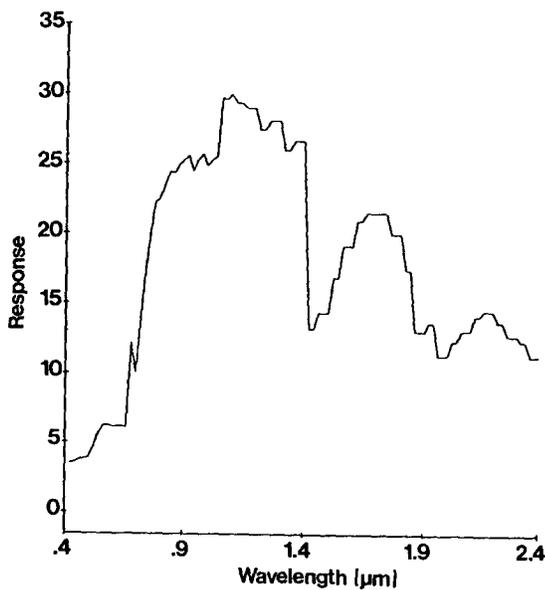
Figure 2. Spectral Response Function for
Mature Wheat Collected on August 4, 1977 Over
Williams County, North Dakota
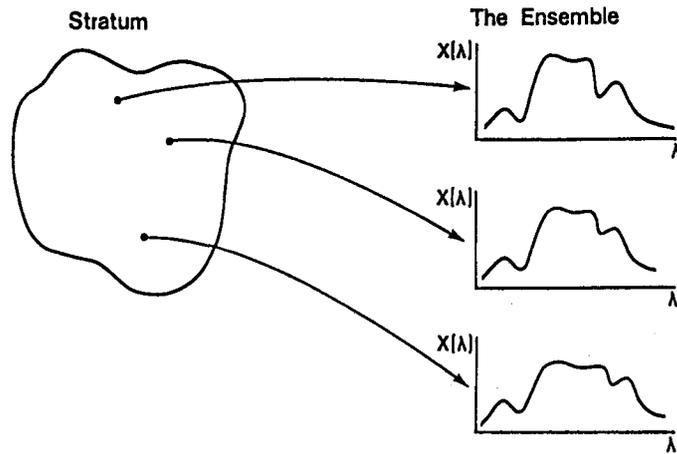
Figure 3. Realization of a Stratum as the
Ensemble of Spectral Sample Functions

$$\hat{x}[\lambda] - X_1\Phi_1[\lambda] + X_2\Phi_2[\lambda] + \ldots \\ + X_N\Phi_N[\lambda]$$
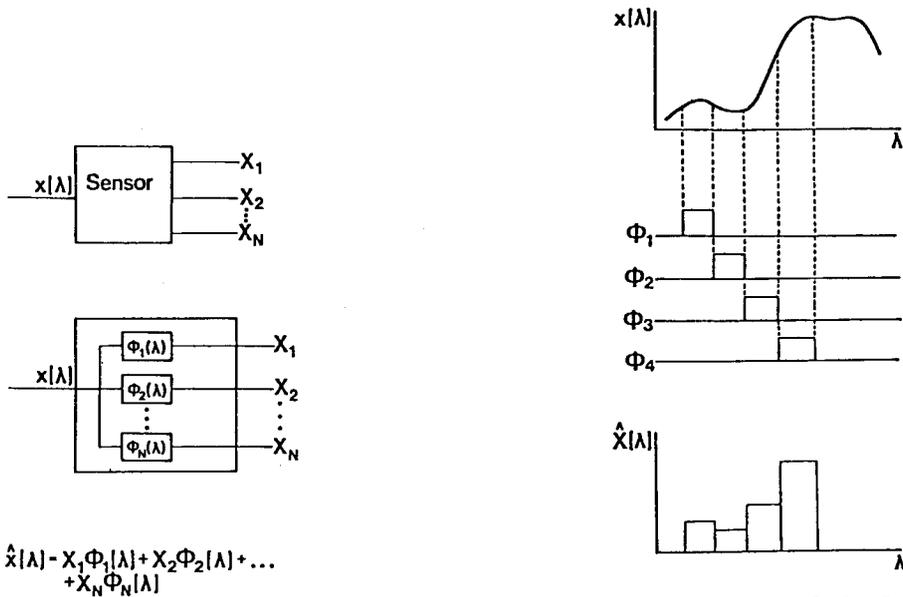
Figure 4. Sensor System Model

Figure 5. Approximation of the Spectral Response
Function by a Set of Four Basis Functions
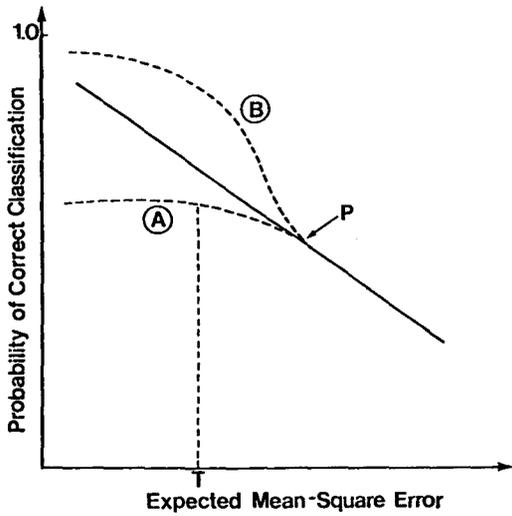
Figure 6. Probability of Correct
Classification as a Function of
Expected Mean-Square Representation
Error
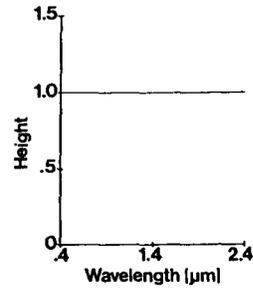
Figure 7. Weight Function Number 1
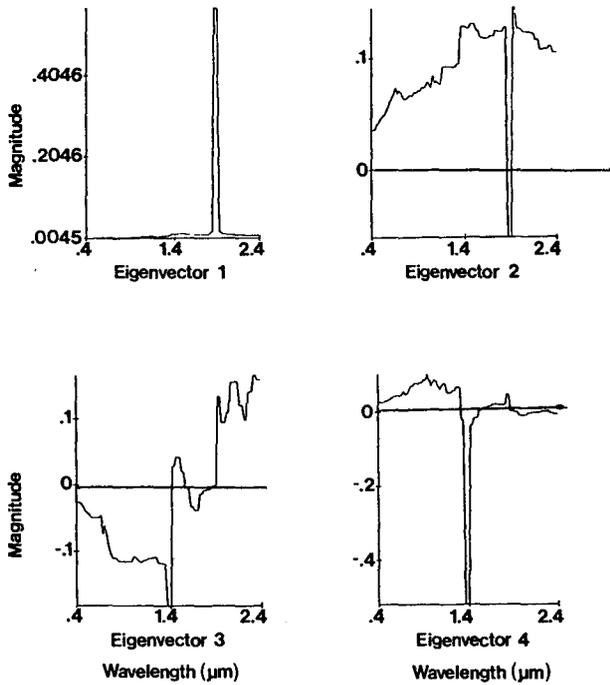
Figure 8. First Four Optimum Basis Functions
Using Weight Function Number 1 Over Williams
County, May 8, 1977 Data
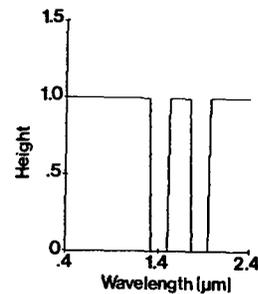
Figure 9. Weight Function Number 2
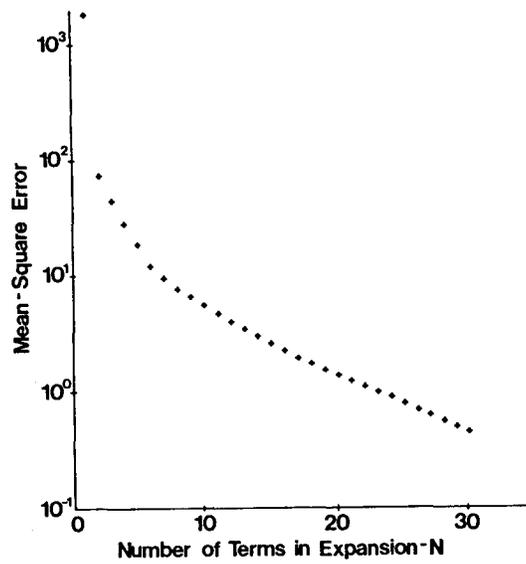
Figure 10. Expected Mean-Square Error as a
Function of the Number of Terms in the
Karhunen-Loeve Expansion for Williams
County, June 29, 1977
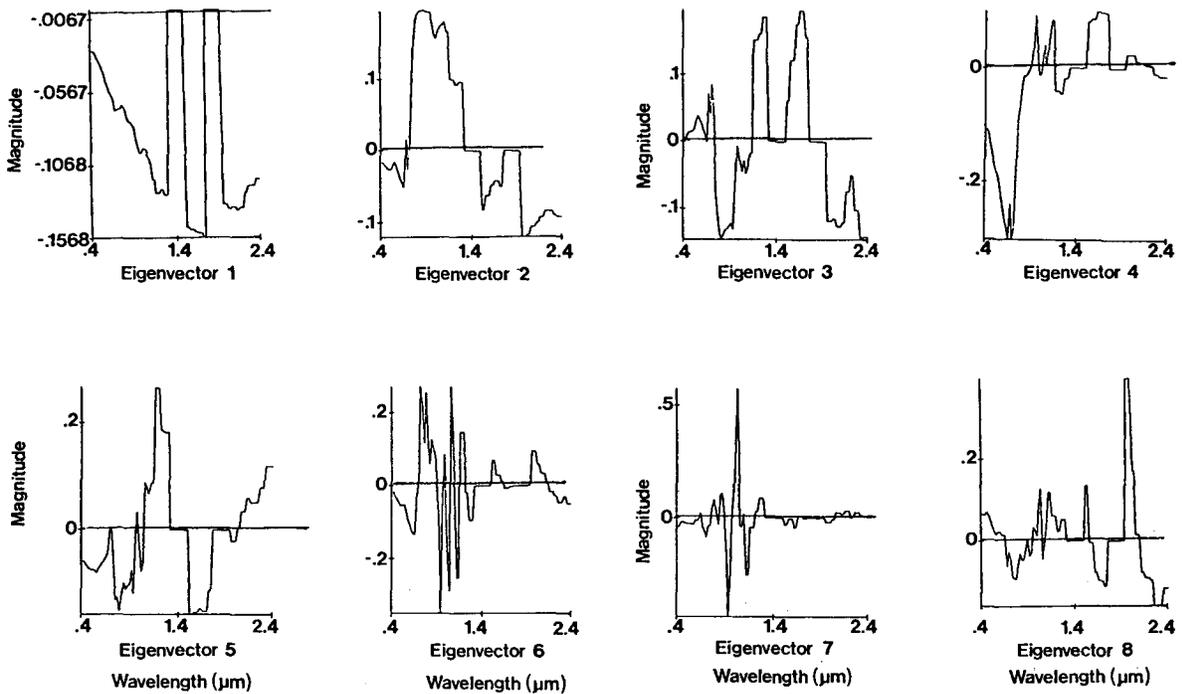
Figure 11. First Eight Eigenvectors for Williams County, June 29, 1977
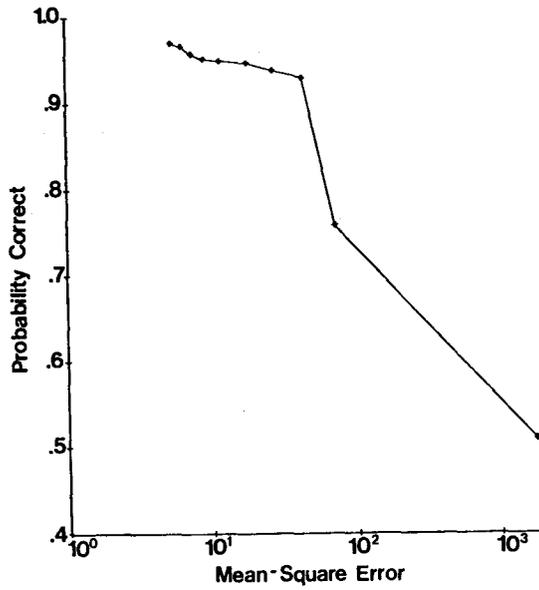
Figure 12. Estimate of Probability of Correct
Classification vs Expected Mean-Square Error for
Williams County, June 29, 1977
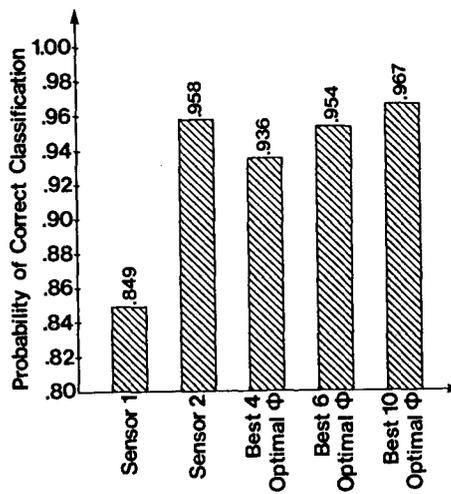
Figure 13. Comparison of Probability of
Correct Classification for Several Sensors
for Williams County, North Dakota, June 29, 1977