END
DATE
FILMED
3 81
DTIC

1.0

1.1

1.25 1.4 1.6

2.8 2.5

3.2 2.2

3.6

4.0 2.0

1.8

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFGL-TR-80-0175

PERFORMANCE OF THE DISCRETE FOURIER TRANSFORM SATELLITE
IMAGERY CLASSIFICATION TECHNIQUE

Robert P. d'Entremont

Systems and Applied Sciences Corporation
6811 Kenilworth Avenue
Riverdale, Maryland 20840

17 June 1980

Scientific Report No. 3

SCIENTIFIC-3

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
AFGL-TR-80-0175	AD-A095364	

4. TITLE (and Subtitle)	5. TYPE OF REPORT & PERIOD COVERED
PERFORMANCE OF THE DISCRETE FOURIER TRANSFORM SATELLITE IMAGERY CLASSIFICATION TECHNIQUE	Scientific Report No. 3
	6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)	8. CONTRACT OR GRANT NUMBER(s)
Robert P. d'Entremont	F19628-79-C-0033

9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
Systems and Applied Sciences Corporation 6811 Kenilworth Avenue Riverdale, Maryland 20840	62101F 667000AA

11. CONTROLLING OFFICE NAME AND ADDRESS	12. REPORT DATE
Air Force Geophysics Laboratory Hanscom AFB, MA 01731 Manager/ Charles Burger/LY	17 June 1980
	13. NUMBER OF PAGES

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report)
	UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Automated satellite imagery classification
Bayes decision theory   Pattern recognition
Cloud classification   Nephanalysis
Cloud detection   Weather satellites
Discrete Fourier transform

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The discrete Fourier transform (DFT) automated satellite imagery classification technique is designed to detect and identify cloud features from 25 x 25 nautical mile (nm) Defense Meteorological Satellite Program (DMSP) visible and infrared imagery samples. The DFT classifier performs two basic steps: feature extraction and sample identification. The feature extraction technique reduces and compresses the original visible and infrared data of an imagery sample in order to retain only the information needed to determine

DD FORM 1473
1 JAN 73

393816

the cloud type of that sample. A multivariate normal density discriminant function then uses this information to identify the cloud type of the image sample from nine possible classes. The results obtained through this experiment show that the classifier can provide accurate and reliable cloud type classifications of satellite imagery samples. This report presents a detailed mathematical description of the DFT spectral classifier technique and offers some ideas for future modifications to the classifier.

Accession For
NTIS GRA&I
DTIC TAB
Unannounced
Justification

By
Distribution/
Availability Codes
Dist | Avail and/or
Special

TABLE OF CONTENTS

1. INTRODUCTION

Large amounts of global cloud cover data in the form of visible and infrared satellite imagery are gathered daily by ground stations throughout the world. These data are readily accessible and available for exploitation. Nevertheless, the extraction of useful information (cloud type, fractional cover, precipitation amounts) from satellite data is not keeping pace with the rate and volume at which these data are collected.

Although valuable when practiced by a skilled analyst, current subjective satellite imagery cloud classification techniques are time-consuming and costly. Hard copies are expensive to produce and distribute on a large-scale basis. Accurate human analysis of all available data is not feasible, and such analyses may lack consistency of interpretation among photoanalysts.

The abundance of satellite data collected by Air Force Global Weather Central (AFGWC) has sparked interest in the development of efficient real-time imagery processing methods which provide meteorological information quickly and accurately. AFGWC currently utilizes automated cloud analysis and classification procedures which determine 3-dimensional global cloud cover distributions using a 25 nautical mile (nm) grid spacing. This process, called 3D Nephanalysis (3DNEPH), uses imagery data from the Defense Meteorological Satellite Program (DMSP) along with surface, radiosonde, and aircraft data. DMSP satellites are equipped with very high resolution or fine mode (1/3 nm) visible and infrared sensors. Nonetheless, satellite tape recorders and AFGWC's computer limitations restrict 3DNEPH satellite input to smoothed (3 nm) imagery for global applications. A decrease in the quantity of input inevitably decreases the quality of output, especially for areas where conventional data are sparse or not available.

Hence there is a need for improving real-time satellite imagery processing capabilities. The incorporation of fine

mode satellite imagery is fundamental for fulfilling the need, even though significantly larger amounts of data will have to be processed. Fast and reliable analytical processing techniques are reasonable and wise approaches to utilizing fine mode imagery data. Techniques must be able to reduce large amounts of satellite imagery data while retaining as much information as possible about any clouds in the imagery.

Discrete Fourier transform power spectrum analysis is the basis of one such technique which has been studied and is described in this report. An algorithm was developed as an alternative to that part of 3DNEPH which determines cloud types. For this study the algorithm was updated and modified for use with 2/3 nm DMSP visible and infrared data. It classifies a 25x25 nm imagery sample as one of nine possible cloud types in less than one second on the Air Force Geophysics Laboratory (AFGL) CDC 6600 computer.

This report presents a detailed mathematical, analytical, and computational description of this cloud classification technique. It also discusses the significance of its results, and offers some ideas for further refining and improvement of the algorithm.

2. DATA

The spectral analysis routine uses visible and infrared data from a DMSP satellite in a 450 nm sun-synchronous orbit. In particular, imagery from several Mediterranean-Middle East orbits was studied. Figures 1a and 1b show samples of visible and infrared imagery from orbit 7399+9 of DMSP vehicle F-1. The Sinai peninsula (A) is located in the lower center of each image. To the west of the Red Sea and the Gulf of Suez, the Nile River can be seen.

Each data value, or pixel, represents the brightness or temperature of a 2/3 nm region of the earth. 2/3 nm DMSP data are unique to this study. DMSP satellites are equipped with 1/3 nm sensors. At AFGWC, the brightness value at

6

Figure 1a. Visible imagery from DMSP orbit #7399+9 (February 1978), showing only every fifth picture element of every fifth line of 2/3 nm imagery. ▲ marks Sinai Peninsula.

Figure 1b. Infrared imagery covering the same area as Figure 1a. Cold temperatures are represented by bright tones and warm temperatures by dark.

8

each point of the 2/3 nm imagery was obtained by averaging four pixels taken from the original 1/3 nm data.

The size of a sample that is classified by this algorithm is 37x37 pixels. It represents a 25x25 nm region of the earth, which is 3DNEPH's horizontal grid spacing, since each pixel represents 2/3 nm.

The Man-computer Interactive Data Access System (McIDAS) at AFGL was used to display the DMSP imagery and store its digital values. McIDAS is capable of storing 24 500 x 672-pixel arrays per user terminal. Thus the data for up to 12 visible and 12 infrared images, each one covering a 333 x 448 nm area of the earth, can be readily accessed. (See Figures 2a and 2b.) The invoking of a McIDAS save command enables a user to search interactively for interesting cloud samples within a displayed DMSP image. Once such a sample is found, the user subjectively evaluates the cloud type in the image sample, along with its background, for verification. This information, in addition to the sample's visible and infrared satellite data, is then written on a sample save tape. This process is continued until the user feels an adequate cloud imagery sample set for use by the automated classifier has been collected.

Figures 2a and 2b show the visible and infrared DMSP imagery of an area of northern Egypt and the northwestern region of the Sinai. The Nile delta is clearly visible in the center of Figure 2a. The darker area in the lower center of the image is El Faiyum, a topographical depression irrigated by the waters of the Nile. The Gulf of Suez is in the lower right. A line of cumuliform clouds lies to the west of the Nile (left-center of the image) covered by a thin cirrostratus shield, hardly detectable in the visible but highly detectable in the infrared. A 25 x 25 nm cloud imagery sample used by the spectral analysis routine covers $\frac{1}{245}$ the size of the area covered by Figures 2a and 2b.

Figure 2a. Visible imagery from DMSP orbit #7399+9 (February 1978). This image is original 2/3 nm resolution. Area shown is the Nile delta.
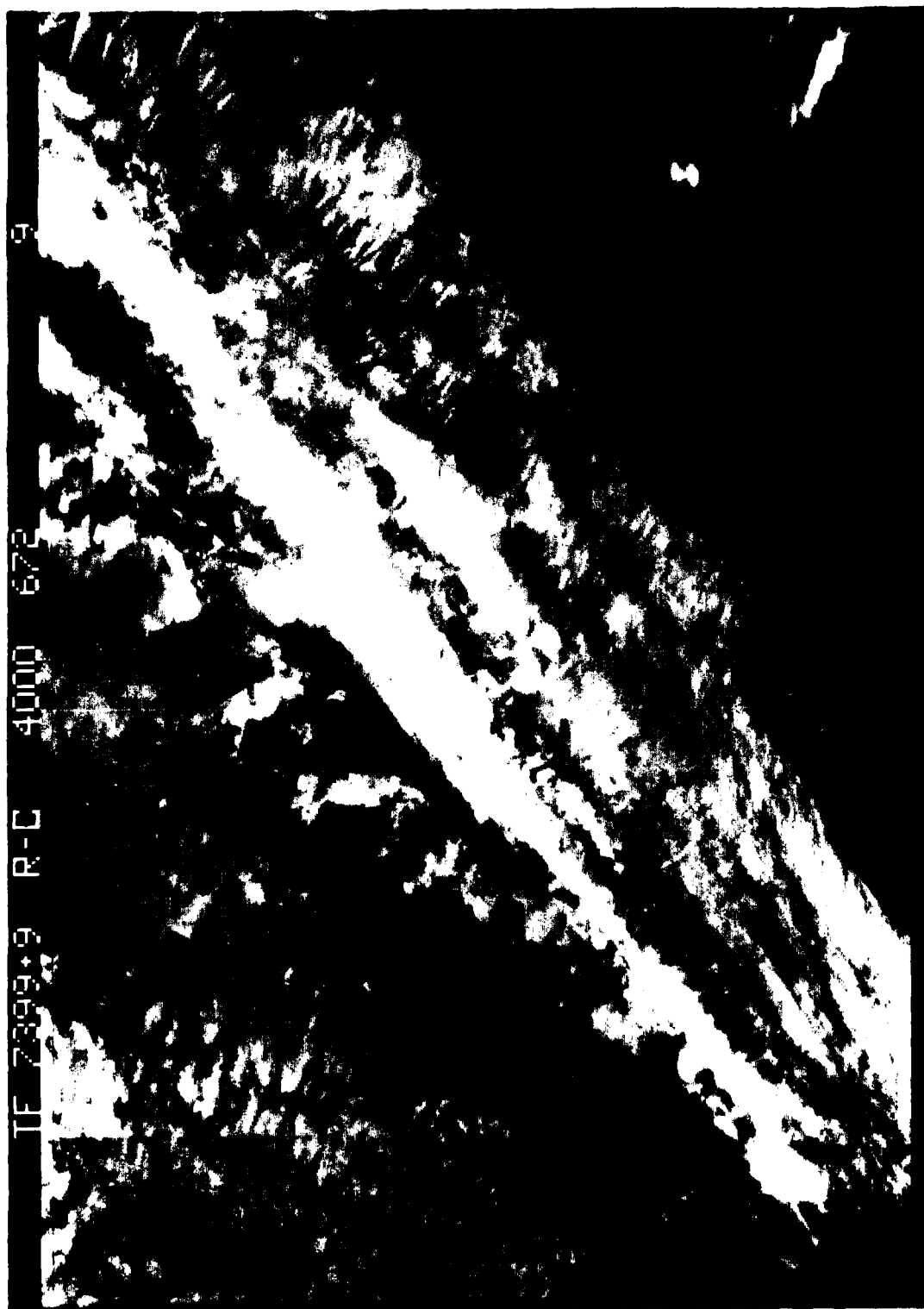
Figure 2b. Infrared imagery covering the same area as Figure 2a. Cold temperatures are represented by bright tones and warm temperatures by dark.

11

3. DESCRIPTION OF THE CLASSIFIER

The spectral classifier technique consists of two main parts: the discrete Fourier transform and a Bayes decision rule. The discrete Fourier transform of a cloud imagery sample is first computed, since the transform contains information about the shapes and sizes of clouds present in the imagery field. The transform is then reduced in size to a power spectrum, which is a vector whose components are normalized averages of the coefficients found in the transform. Then a Bayes discriminant function decides the cloud type of the satellite imagery sample using the information contained in the power spectrum. The discriminant function is derived from the multivariate normal density function.

In the remainder of the section, a mathematical formalization of this technique is developed in order to provide a better understanding of the spectral classifier. Symbols are defined in the Glossary of Terms, page 74.

The continuous Fourier transform has been a standard analytical tool for solving many types of pattern recognition problems. However, until recently, applications of the discrete Fourier transform (DFT) have been limited because of the large amounts of computational time required to calculate it. With the development of the fast Fourier transform, an algorithm which computes the DFT quickly and efficiently, the DFT's use in problem solving has greatly increased.

3.1 The Discrete Fourier Transform

A real-valued, continuous function $F(x,y)$ of two variables defined on the region $\mathscr{R}$ where $0 < x \leq M$, $0 < y \leq N$, M and N positive integers, can be expressed as a unique linear combination of products of sines and cosines:

$$\mathbf{F}(x,y) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \left\{ A_{m,n} \cos \frac{2\pi mx}{M} \cos \frac{2\pi ny}{N} + B_{m,n} \cos \frac{2\pi mx}{M} \right.$$

$$\sin \frac{2\pi ny}{N} + C_{m,n} \sin \frac{2\pi mx}{M} \cos \frac{2\pi ny}{N} + D_{m,n} \sin \frac{2\pi mx}{M}$$

$$\left. \sin \frac{2\pi ny}{N} \right\}, \tag{1}$$

for $(x,y) \in \mathcal{R}$, and where $A_{m,n}$, $B_{m,n}$, $C_{m,n}$, and $D_{m,n} \in \mathbf{R}$. This is called the double Fourier series expansion of $\mathbf{F}$ on $\mathcal{R}$. This representation extends $\mathbf{F}$ periodically for all $(x,y) \in \mathbf{R}^2$. An alternate, shorter form is:

$$\mathbf{F}(x,y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \psi_{m,n} e^{2\pi\sqrt{-1}(\frac{mx}{M} + \frac{ny}{N})}, \quad \psi_{m,n} \in \mathbf{C} \tag{2}$$

where
$$\psi_{m,n} = \frac{1}{MN} \int_0^M \int_0^N \mathbf{F}(x,y) e^{-2\pi\sqrt{-1}(\frac{mx}{M} + \frac{ny}{N})} dy\, dx, \quad \forall m,n \in \mathbf{Z}. \tag{3}$$

Note that $\psi_{0,0}$ is the average value of $\mathbf{F}$ over $\mathcal{R}$. Equation (3) is known as the <u>double</u> <u>Fourier</u> <u>transform</u> of $\mathbf{F}$ on $\mathcal{R}$.

There are infinitely many harmonics in $\mathbf{F}$, as can be seen by looking at the upper and lower bounds of the $\sum$'s in (1) and (2). Infinitely many coefficients $\psi_{m,n}$ can be solved for by (3) since $\mathbf{F}$ is continuous on $\mathcal{R}$. Equations (1) and (2) are exact representations of $\mathbf{F}(x,y)$ on $\mathcal{R}$.

Let $\hat{\mathcal{R}} = \{(i,j) | (i,j) \in \mathcal{R} \text{ and } i,j \in \mathbf{Z}\}$, where $\mathbf{Z}$ is the set of integers. Consider the discrete function

$$\hat{\mathbf{F}}(i,j) = \mathbf{F}(i,j) \quad \forall (i,j) \in \hat{\mathcal{R}}.$$

$\hat{\mathbf{F}}$ can be extended to the entire lattice of points (i,j) on the plane $\mathbf{R}^2$ such that

$$\hat{\mathbf{F}}(i+kM, j+lN) = \hat{\mathbf{F}}(i,j) \text{ for any } k,l \in \mathbf{Z},$$

that is, $\hat{\mathbf{F}}$ is periodic on $\mathbf{Z} \times \mathbf{Z}$. $\hat{\mathbf{F}}$ can be expressed as a finite series in the form:

$$\hat{\mathbf{F}}(i,j) = \sum_{m=0}^{\left[\frac{M}{2}\right]} \sum_{n=0}^{\left[\frac{N}{2}\right]} \left\{ \hat{A}_{m,n} \cos \frac{2\pi mi}{M} \cos \frac{2\pi nj}{N} + \hat{B}_{m,n} \cos \frac{2\pi mi}{M} \right.$$

$$\sin \frac{2\pi nj}{N} + \hat{C}_{m,n} \sin \frac{2\pi mi}{M} \cos \frac{2\pi nj}{N} + \hat{D}_{m,n} \sin \frac{2\pi mi}{M}$$

$$\left. \sin \frac{2\pi nj}{N} \right\}, \tag{4}$$

where $\widehat{A}_{m,n}$, $\widehat{B}_{m,n}$, $\widehat{C}_{m,n}$, $\widehat{D}_{m,n} \in \mathbf{R}$. This is called the discrete double Fourier series expansion of $\widehat{F}$ on $\widehat{\mathcal{R}}$. Its shorter, complex form is:

$$\widehat{F}(i,j) = \sum_{m=\left[\frac{-M}{2}\right]+1}^{\left[\frac{M}{2}\right]} \sum_{n=\left[\frac{-N}{2}\right]+1}^{\left[\frac{N}{2}\right]} \hat{\psi}_{m,n} e^{2\pi\sqrt{-1}\left(\frac{mi}{M}+\frac{nj}{N}\right)}, \quad \hat{\psi}_{m,n} \in \mathbf{C}, \quad (5)$$

where $[x]$ is the greatest integer less than or equal to x, x positive, or the smallest integer greater than or equal to x, x negative.[1] The $\hat{\psi}_{m,n}$ are given by equation (6):

$$\hat{\psi}_{m,n} = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \widehat{F}(i,j) e^{-2\pi\sqrt{-1}\left(\frac{mi}{M}+\frac{nj}{N}\right)}, \quad m,n \in \mathbf{Z}$$

such that $\left[\frac{-M}{2}\right] < m \leq \left[\frac{M}{2}\right]$, and $\left[\frac{-N}{2}\right] < n \leq \left[\frac{N}{2}\right]$. $\qquad\qquad$ (6)

Note that $\hat{\psi}_{0,0}$ is the average value of $\widehat{F}$ over $\widehat{\mathcal{R}}$. Equation (6) is known as the <u>discrete double Fourier transform</u> of $\widehat{F}$ on $\widehat{\mathcal{R}}$.

The coefficients in any Fourier series expansion of a discrete function $\widehat{F}$ are the amplitudes of the trigonometric functions they precede.[2] Each coefficient in this expansion contains information pertaining to the size of a discrete harmonic that is present in $\widehat{F}$.

Since it is discrete, $\widehat{F}$ contains only a finite number of harmonics. There is a limit to the number of unique coefficients $\hat{\psi}_{m,n}$ that can be determined. This statement is proved in Appendix B.

In a discrete sense, the smallest resolvable nontrivial harmonic (wave) in either the x- or y-direction needs at least two points to establish itself, so that the largest wavenumber $\widehat{F}$ can possibly contain in the x-direction, given M points in the x-direction, is $\left[\frac{M}{2}\right]$, and in the y-direction, given N points in the y-direction, is $\left[\frac{N}{2}\right]$. It is impossible

[1]See Appendix A for the relationship between (5) and (4).
[2]This is easy to see in equation (4). In equation (5), remember $e^{i\theta} = \cos\theta + i\sin\theta$, $i = \sqrt{-1}$.

... ... harmonic to ... ... thin ... ...
... ... $\mathscr{R}$ of $\hat{\mathbf{F}}$ more than $\left[\frac{M}{2}\right]$ ... ...
and more than $\left[\frac{N}{2}\right]$ times in the y-direction.

Thus, in equation (4), there are $\frac{M}{2} \cdot \frac{N}{2}$ $\hat{A}_{m,n}$'s, $\hat{B}_{m,n}$'s,
$\hat{C}_{m,n}$'s, and $\hat{D}_{m,n}$'s, or $4 \cdot \frac{MN}{4} = MN$ distinct, nontrivial
solvable Fourier coefficients. Likewise, in (5), there are
also MN distinct $\hat{\psi}_{m,n}$'s. In general only MN coefficients can
be computed when $\hat{\mathbf{F}}$ is represented by MN values.[3]

The numbers $\left[\frac{M}{2}\right]$ and $\left[\frac{N}{2}\right]$ are called the **Nyquist** wave-
numbers of $\hat{\mathbf{F}}$ on $\mathscr{R}$ in the x- and in the y-directions,
respectively.

3.1.1 The Significance of Fourier Coefficients

Each coefficient in the Fourier expansion of a discrete
function contains information pertaining to the amplitude of
a discrete harmonic that is present in that function. For
example, if there exists no wave in $\hat{\mathbf{F}}$ with wavenumber
$m = i$, $\left[\frac{-M}{2}\right] < i \le \left[\frac{M}{2}\right]$ in the x-direction and $n = j$, $\left[\frac{-N}{2}\right] < j$
$\le \left[\frac{N}{2}\right]$ in the y-direction, then the Fourier coefficients in
(4) and in (5) with the subscript (wavenumber) "i,j" are
equal to zero.

The harmonics of a Fourier expansion are generated about
the mean of the function which the expansion represents, as is
shown in Figure 3.

The Fourier series representation of the discrete func-
tion of one variable shown in Figure 3a is (set j, n = 0 in
(4)):
$$\hat{y} = 0.75 - 0.25 \cos \pi x,$$
for x = 1,2,3,...,8. Note how the cosine wave (dashed curve)
of wavenumber 4 and amplitude 0.25 ($\hat{A}_4 = -0.25$) lies about
the line y = 0.75, which is the mean $\hat{A}_0$ of $\hat{y}$.

In two dimensions, the mean of a function is not a line

---

[3]Just as $\hat{\mathbf{F}}$ is periodic, so is its transform. A function
$\hat{\mathbf{F}}(i,j)$ of space is transformed into a function $\hat{\psi}_{m,n}$ of
wavenumber.
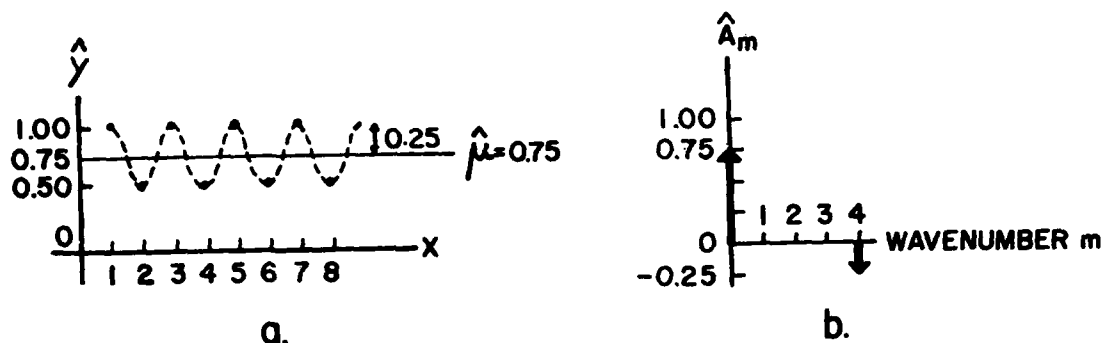
Figure 3. A discrete cosine wave $\hat{y}$ of amplitude 0.25 which lies about its mean 0.75. $\hat{y}$ is assumed periodic outside of the region shown here. Note how the "pattern" established in the first two points ($x = 1$ and $x = 2$) occurs four times over the period of $\hat{y}$, thus making $\hat{A}_4$ nonzero. Figures 3a and 3b are Fourier transform pairs: $\hat{A}_0 = 0.75$, and $\hat{A}_4 = -0.25$.
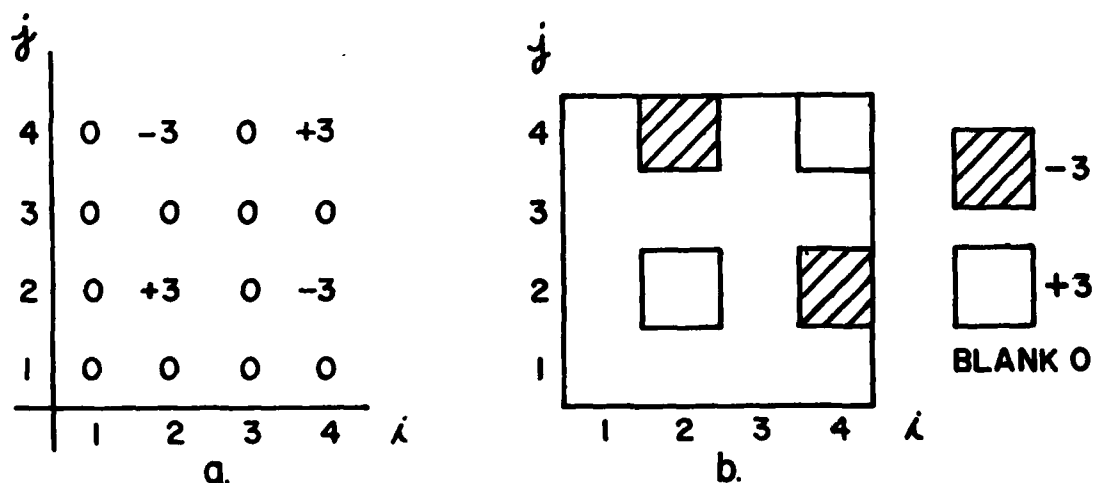
Figure 4. A simple discrete, finite function $\hat{G}(i,j)$ of two variables defined on the region $\mathcal{D} = \{(i,j) \mid 1 \leq i \leq 4,\ 1 \leq j \leq 4,\ i$ and j integers$\}$. $\hat{G}(i,j)$ is assumed periodic outside of $\mathcal{D}$.

but a plane of constant distance $\widehat{A}_{0,0}$ (which is equal to $\hat{\psi}_{0,0}$) from the x-y plane. The harmonics present in a function of two variables take on various values above and below its mean plane.

Consider the discrete function $\widehat{G}$ of two variables and its Fourier coefficients. Let $M = N = 4$, and let $\widehat{G}(i,j)$ be defined as in Figure 4a. Computing the transform of $\widehat{G}$ for the form (4), we find $\widehat{A}_{1,1} = 3$ and all other 15 coefficients are identically zero. This is true because $\widehat{G}$ is a discrete, pure harmonic of wavenumber $m = 1$, $n = 1$. By looking at Figure 4b it is more easily seen that there is in fact a discrete, two-dimensional harmonic in $\widehat{G}$ that occurs exactly once horizontally ($m = 1$) and exactly once vertically ($n = 1$). Instead of defining $\widehat{G}$ as a discrete set of sixteen ordered triples $(i, j, \widehat{G}(i,j))$, a more compact expression in the form (4) would be:

$$\widehat{G}(i,j) = \widehat{A}_{1,1} \cos \frac{2\pi \cdot 1 \cdot i}{4} \cos \frac{2\pi \cdot 1 \cdot j}{4}$$

$$= 3 \cos \frac{\pi i}{2} \cos \frac{\pi j}{2}, \qquad (7)$$

where $i, j \in \mathbf{Z}$ such that $1 \leq i \leq 4$, $1 \leq j \leq 4$.

Now suppose we define a new discrete function $\widehat{H}(i,j)$ as in Figure 5a. Transforming this function we get $\widehat{A}_{2,2} = 3$, and all other 15 coefficients equal zero. $\widehat{H}$ is a pure "cosine-cosine" wave of wavenumber $m = 2$, $n = 2$. As can be seen in Figure 5b, a discrete harmonic occurs twice in the horizontal ($m = 2$) and twice in the vertical ($n = 2$) in the principal part of its domain, and has amplitude 3:

$$\widehat{H}(i,j) = \widehat{A}_{2,2} \cos \frac{2\pi \cdot 2 \cdot i}{4} \cos \frac{2\pi \cdot 2 \cdot j}{4}$$

$$= 3 \cos \pi i \cos \pi j. \qquad (8)$$

Equation (8) is in the form (4); i and j are as defined for $\widehat{G}$.

As was seen in the previous two examples, the Fourier transform of a discrete function of two variables is sensi-
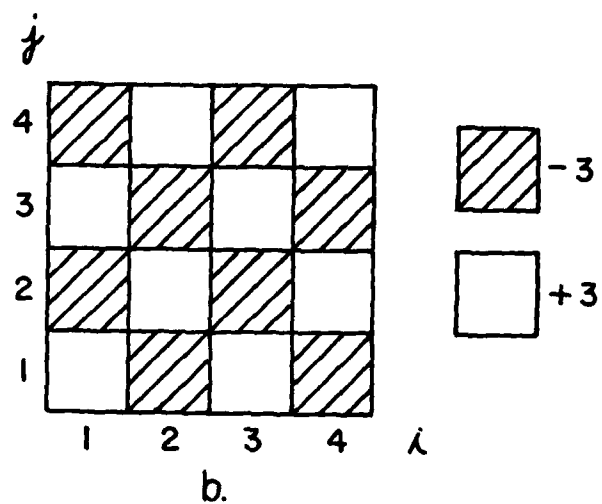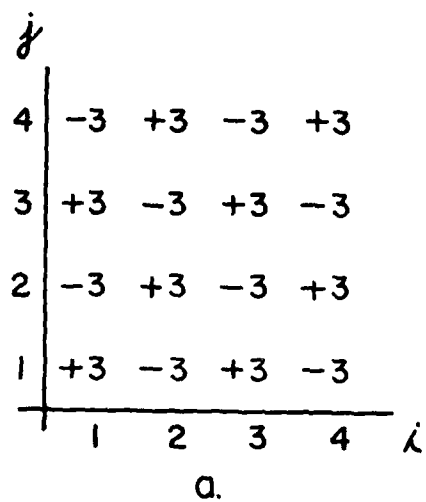
Figure 5. A simple discrete, finite function $\hat{H}(i,j)$ of two
variables defined on the region $\hat{\mathcal{D}}$ as defined in Figure 4.
$\hat{H}(i,j)$ is assumed periodic outside of $\hat{\mathcal{D}}$.

tive to the shapes (wavenumbers) and sizes (amplitudes) of
the harmonics in that function. With this in mind, consider
the function $\hat{\mathbf{F}}$ to be a finite set of visible or infrared
satellite data for a particular region $\hat{\mathscr{R}}^4$ on the earth. $\hat{\mathbf{F}}$
is certainly a discrete function of two variables. As long
as it is assumed periodic outside of $\hat{\mathscr{R}}^5$, its transform can
be computed. The coefficients of this Fourier transform can
provide useful information pertaining to the shapes and sizes
of any clouds present in $\hat{\mathbf{F}}$ over $\hat{\mathscr{R}}$.

This information is the basis of the spectral classi-
fier technique. The DFT of a 37 x 37-pixel array $\hat{\mathbf{F}}$ (25 x
25 nm) containing the satellite-sensed visible or infrared
brightnesses of a cloud sample is computed. The coefficients
thus obtained may then be examined to determine the texture
and patterns of the clouds in the imagery sample.

3.1.2 The Fast Fourier Transform

The DFT of a function can be computed from equations
like (6), but for large M and large N this is a slow opera-
tion. It is then more rapidly calculated by the algorithm
which is known as the fast Fourier transform (FFT).

The FFT is a more efficient, quicker way of computing
the coefficients found in the complex Fourier expansion of
a discrete, finite function. For example, the FFT algorithm
calculates the Fourier transform of a one-dimensional array
of 1,024 values over 200 times faster than the straight-
forward method. Moreover, the output of an FFT is inter-
preted in exactly the same manner as the output of a DFT.

Several restrictions are placed on any data entered
into the FFT. As is the case with the DFT, the number of
coefficients in the transform of a function must equal the
number of values defining that function. The input data

---

[4]In other words, a finite set of latitude-longitude grid points.
[5]We are only interested in $\hat{\mathbf{F}}$ on $\hat{\mathscr{R}}$ and nowhere else, so this
assumption has no adverse consequences.

must represent equally-spaced values of a function in the
space domain, and the output must represent equally-spaced
coefficients of its transform in the wavenumber domain. The
input function and its transform are assumed to be one com-
plete cycle of discrete periodic functions. The coefficients
generated by the FFT are not normalized; for an M x N input
array they appear multiplied by M·N.

In general, these restrictions are minimal and the FFT
is a valuable and powerful computational tool when the DFT
of a large function of two variables must be calculated.

3.1.3 Annular Integration

The transform of a 37 x 37-pixel array $\hat{F}$ contains
$37 \cdot 37 = 1,369$ coefficients. Analyzing each coefficient
separately would be a painstaking, time-consuming approach.
Therefore, the number of these coefficients must somehow be
reduced, while at the same time retaining as much information
about the clouds in the imagery sample $\hat{F}$ as possible. The
data reduction technique used by this classifier is called
annular integration and is described here.

The square root of the sum of the squares of the real
and imaginary parts of a Fourier coefficient $\hat{\psi}_{m,n}$ (6) is
defined as <u>power</u>:

$$P_{m,n} = \sqrt{Re(\hat{\psi}_{m,n})^2 + Im(\hat{\psi}_{m,n})^2} . \qquad (9)$$

Just as a Fourier coefficient is a measure of how far a
harmonic deviates from the mean of a function (amplitude),
so is power, since it is defined in terms of the components
of a Fourier coefficient. The set of powers of all the co-
efficients of the transform of a function $\hat{F}$ is called the
<u>Fourier</u> <u>spectrum</u> of $\hat{F}$.

A power can be calculated for every (m,n) of the wave-
number plane for which there is a $\hat{\psi}_{m,n}$. These terms can be
placed onto the m-n plane by mapping each $P_{m,n}$ to the point
(m,n). Then, the powers can be sorted into annular bands
whose components are the same ρ-distance from the origin, where

$$\rho = \left[\sqrt{m^2 + n^2} + 0.5\right], \quad \text{for } \left[\frac{-M}{2}\right] < m \le \left[\frac{M}{2}\right], \quad \text{and} \quad \left[\frac{-N}{2}\right] < n \le \left[\frac{N}{2}\right], \quad (10)$$

and where [ ] is as previously defined.

Given an M x N array of coefficients, the largest number m can be is $\left[\frac{M}{2}\right]$, and the largest number n can be is $\left[\frac{N}{2}\right]$. Therefore the annular band with the largest "radius" ρ_{max} must be (from (10)):

$$\rho_{max} = \left[\sqrt{\left[\frac{M}{2}\right]^2 + \left[\frac{N}{2}\right]^2} + 0.5\right]. \quad (11)$$

It is obvious that

$$\rho_{min} = \left[\sqrt{0^2 + 0^2} + 0.5\right] = 0. \quad (12)$$

For each ρ, every power term in annular band ρ is squared, and these squares are summed together. Finally, they are averaged, this average's square root is taken, and then it is normalized. This number is called a _normalized average amplitude_ (NAA), and is given by:

$$NAA_\rho = \frac{1}{M \cdot N}\sqrt{\frac{1}{NT(\rho)}\sum_\rho P_{m,n}^{\,2}}, \quad \rho = 0,1,2,\ldots,\rho_{max}, \quad (13)$$

where $NT(\rho)$ is the number of power terms in annular band ρ, $\sum_\rho P_{m,n}^{\,2}$ is the sum of the squares of the powers in annular band ρ, and $\frac{1}{MN}$ is the normalization factor. (This factor is necessary only when the transform of a function $\hat{F}$ has been computed by the FFT, as in this study. This is because the coefficients from an FFT appear multiplied by M·N, that is, they appear unnormalized. Whenever equation (6) is used to compute a discrete transform, the coefficients are already normalized by $\frac{1}{MN}$, thus eliminating the need for normalization by (13).) Therefore, an array of $\rho_{max}+1$ normalized average amplitudes NAA_ρ can be calculated from the discrete Fourier transform of a cloud sample. When these NAA's are arranged in order of increasing ρ, the resulting vector is called a _power spectrum_.

For a 37 x 37 transform array the annular integration technique reduces 1,369 Fourier coefficients to a

$$\rho_{max} + 1 = \left[\sqrt{\left[\frac{37}{2}\right]^2 + \left[\frac{37}{2}\right]^2} + 0.5\right] + 1 = 26\text{-component power}$$

spectrum (a reduction of over 52:1).

In order to better understand the concept of annular integration of power, consider this example. Let $M = N = 8$. The number of distinct, concentric annular bands generated by (10) is

$$\left[\sqrt{\left[\frac{8}{2}\right]^2 + \left[\frac{8}{2}\right]^2} + 0.5\right] + 1 = \left[\sqrt{32} + 0.5\right] + 1 = 7.$$

Figure 6 shows the seven annular bands of an 8 x 8 discrete Fourier transform.

NAA_ρ can be thought of as a measure of the amplitude of the harmonics in $\hat{F}$ with wavenumber ρ in one direction. All coefficients whose powers are used to calculate NAA_ρ contain at least one subscript m or n which is equal to or is "nearly equal to" ρ. Each NAA_ρ is a weighted average of the amplitudes of any two-dimensional waveforms of wavelength either "approximately" $\frac{M}{\rho}$ pixels in the x-direction or "approximately" $\frac{N}{\rho}$ pixels in the y-direction, given that $\hat{F}$ is defined for all points (i,j) on an M x N grid.

For a 37 x 37 array of satellite data values which represent a 25 x 25 nm region of the earth, the wavelength of a harmonic in an imagery field $\hat{F}$ with wavenumber ρ in either the x-direction or the y-direction is inversely proportional to the wavenumber:

$$\begin{aligned}
\lambda &= \left[\frac{37}{\rho}\right] \text{ pixels/cycle} \\
&= \left(\left[\frac{37}{\rho}\right] \text{ pixels/cycle}\right) \cdot \left(\frac{2}{3} \text{ nm/pixel}\right) \\
&\simeq \frac{25}{\rho} \text{ nm/cycle}.
\end{aligned} \tag{14}$$

For this study, therefore, NAA_ρ is the average amplitude of all two-dimensional waves (cloud patterns) whose wavelengths are $\frac{25}{\rho}$ nm in either the x-direction or the y-direction of the cloud imagery sample $\hat{F}$. Figure 7 shows wavelengths as a function of ρ.

As an example, it is easy to compute the power spectra of $\hat{G}$ (i,j) (7) and of $\hat{H}$ (i,j) (8):

$\hat{G}$	$\hat{H}$
$NAA_0 = 0$	$NAA_0 = 0$
$NAA_1 = \frac{3\sqrt{2}}{8}$	$NAA_1 = 0$
$NAA_2 = 0$	$NAA_2 = 0$
$NAA_3 = 0$	$NAA_3 = \frac{3}{4}$

Note that only one NAA_ρ is nonzero in each case. There is
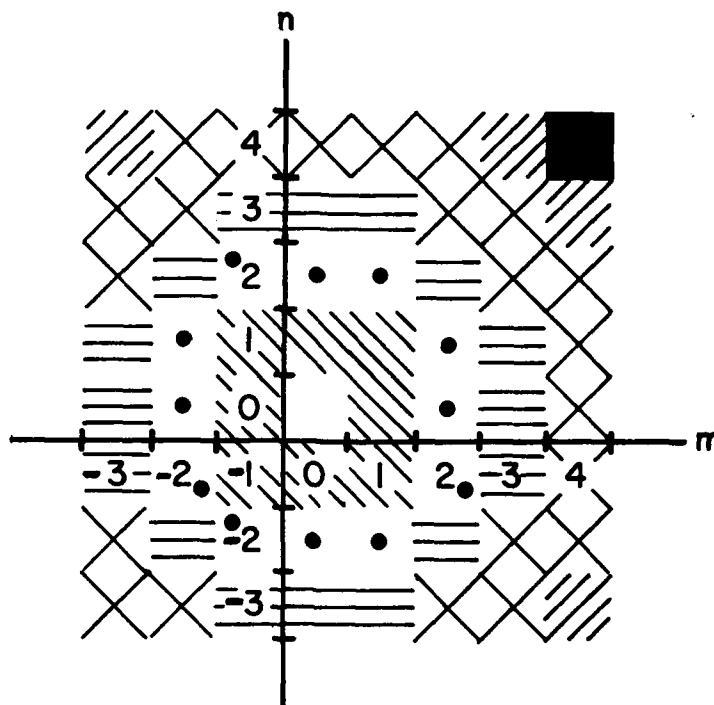
Figure 6. The seven annular bands of the transform of an 8 x 8 array as generated by (10). Each "block" on the m-n plane represents the power $P_{m,n}$ of a Fourier coefficient $\hat{\psi}_{m,n}$ (see (9)).

23
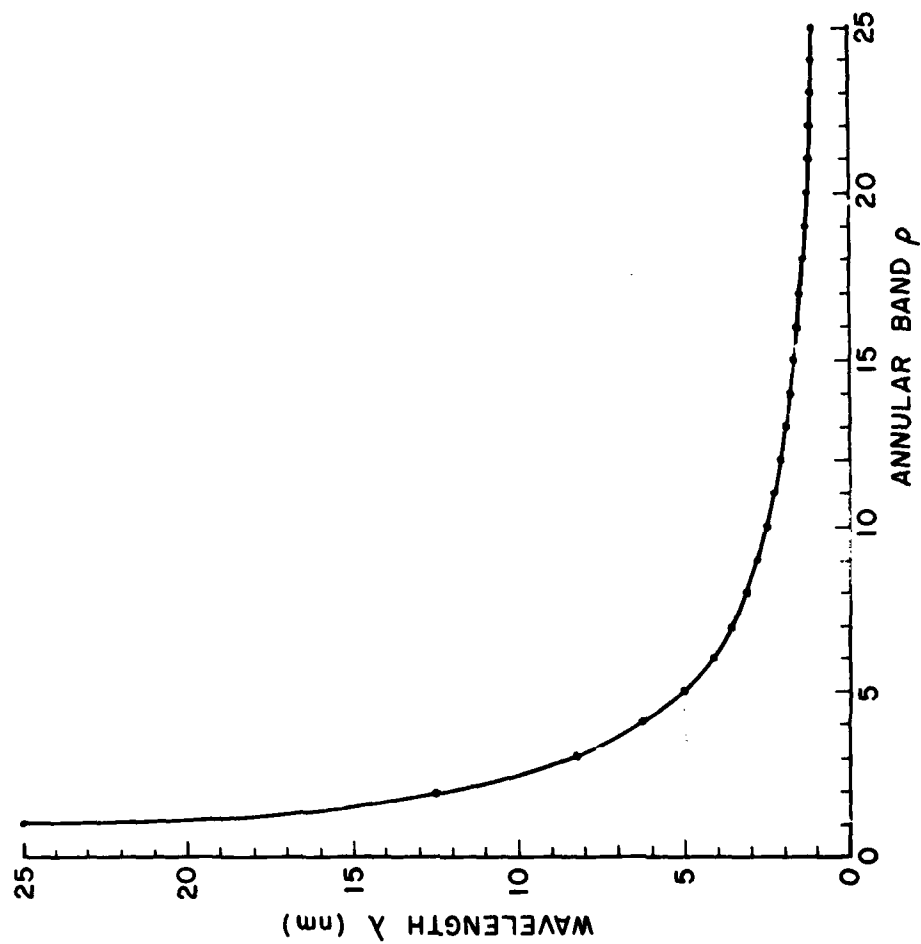
Figure 7. Wavelength as a function of annular band number ρ.

24

enough of a difference between these two spectra to be able to tell which spectrum is $\hat{\mathbf{G}}$'s and which spectrum is $\hat{\mathbf{H}}$'s. Such differences will be used to determine a sample's cloud type given its spectrum.

It is interesting to note that annular band 0 contains the power of just one coefficient, $\hat{\psi}_{0,0}$. From the FFT, $\hat{\psi}_{0,0}$ is the unnormalized average of $\hat{\mathbf{F}}$ over $\hat{\mathscr{R}}$. Since $\hat{\mathbf{F}}$ is real-valued, $\mathrm{Im}(\hat{\psi}_{0,0}) = 0$. Letting $\rho = 0$, from (13) we have:

$$\begin{aligned}
\mathrm{NAA}_0 &= \frac{1}{M \cdot N} \sqrt{\frac{1}{1} \sum_0 P_{m,n}{}^2} \\
&= \frac{1}{M \cdot N} \sqrt{P_{0,0}{}^2} \\
&= \frac{1}{M \cdot N} \sqrt{(\text{unnormalized average})^2} \\
&= \frac{\text{unnormalized average}}{\text{normalization factor}} \\
&= \text{mean of } \hat{\mathbf{F}}.
\end{aligned} \qquad (15)$$

One important property of the annular integration technique is that it is not sensitive to directionality. Any finite discrete periodic function $\hat{\mathbf{F}}(i,j)$ on $\hat{\mathscr{R}}$ rotated any integer multiple of $\frac{\pi}{2}$ radians[6] will have the same power spectrum it "originally" had, even though most of the Fourier coefficients change. The coefficients in a Fourier transform are sensitive to the absolute directionality (orientation) of the harmonics in a function, whereas a power spectrum is not. In other words, a power spectrum retains information relative to the shapes and sizes of clouds in an imagery sample $\hat{\mathbf{F}}$ but loses the sense of their directionality. This is an advantageous property for a power spectrum to have, since the north-south, east-west alignment of clouds should not affect their classification. It is for this reason that the spectral classifier uses the imagery sample's power spectrum to identify its cloud type.

---

[6]One need only be concerned with 90° rotations, since any other rotation of $\hat{\mathbf{F}}$ on $\hat{\mathscr{R}}$ would result in $\hat{\mathbf{F}}$'s not being defined for all grid points in $\hat{\mathscr{R}}$. Remember $\hat{\mathbf{F}}$ may change but $\hat{\mathscr{R}}$ remains fixed.

Figure 8 is an example of the rotation property. Function (a) is rotated 90° about the point $(1,1)$ to obtain function (b). The average value $\hat{\psi}_{0,0}$ of each function is 2, but in (a) the only other nonzero coefficient is $\hat{\psi}_{4,0} = 1$, whereas in (b) the only other nonzero coefficient is $\hat{\psi}_{0,4} = 1$. It is easily seen that $\hat{\psi}_{4,0}$ (from (a)) and $\hat{\psi}_{0,4}$ (from (b)) are used to calculate each function's respective NAA_4, since (from (10)):

$$\left[\sqrt{4^2 + 0^2} + 0.5\right] = \left[\sqrt{0^2 + 4^2} + 0.5\right] = 4.$$

The patterns in functions 8(a) and 8(b) are the same except for their orientation. Their power spectra are equal, even though the functions and their transforms are different.

An infinite number of test patterns could have the power spectrum given by Figure 8. In addition to the cases given, there could be variations in both the i and j directions since any of the 22 Fourier coefficients in $\rho = 4$ could be nonzero so long as the sum of the squares of their powers is equal to 1.

Annular integration is an effective data reduction technique and a useful tool for objectively extracting cloud size information from an imagery sample $\hat{F}$. As with any data reduction technique, some information is lost. This technique loses the sense of direction of the cloud elements. It can also confuse highly directional cloud features, e.g., cloud bands, with more complex scenes having variation in all directions, e.g., cumulus clouds. The losses are, however, adequately compensated by the convenience and speed of working with a one-dimensional power spectrum. Using AFGL's CDC 6600 computer, the fast Fourier transform algorithm calculates the discrete Fourier transform of a 37 x 37 input array $\hat{F}$ (1,369 numbers) in 7/10 sec, and the calculation and annular sorting and summing of the powers of these coefficients are done in 3/100 sec. Such computational speed can allow for quick cloud sample identification.

$\hat{\psi}_{0,0} = 2$

$\hat{\psi}_{4,0} = 1 + 0\sqrt{-1}$

AND ALL OTHER COEFF. = 0

$NAA_0 = 2$

$NAA_1 = 0$

$NAA_2 = 0$

$NAA_3 = 0$

$NAA_4 = \dfrac{1}{\sqrt{22}}$

$NAA_5 = 0$

$NAA_6 = 0$

$\hat{\psi}_{0,0} = 2$

$\hat{\psi}_{0,4} = 1 + 0\sqrt{-1}$

AND ALL OTHER COEFF. = 0

$NAA_0 = 2$

$NAA_1 = 0$

$NAA_2 = 0$

$NAA_3 = 0$

$NAA_4 = \dfrac{1}{\sqrt{22}}$

$NAA_5 = 0$

$NAA_6 = 0$
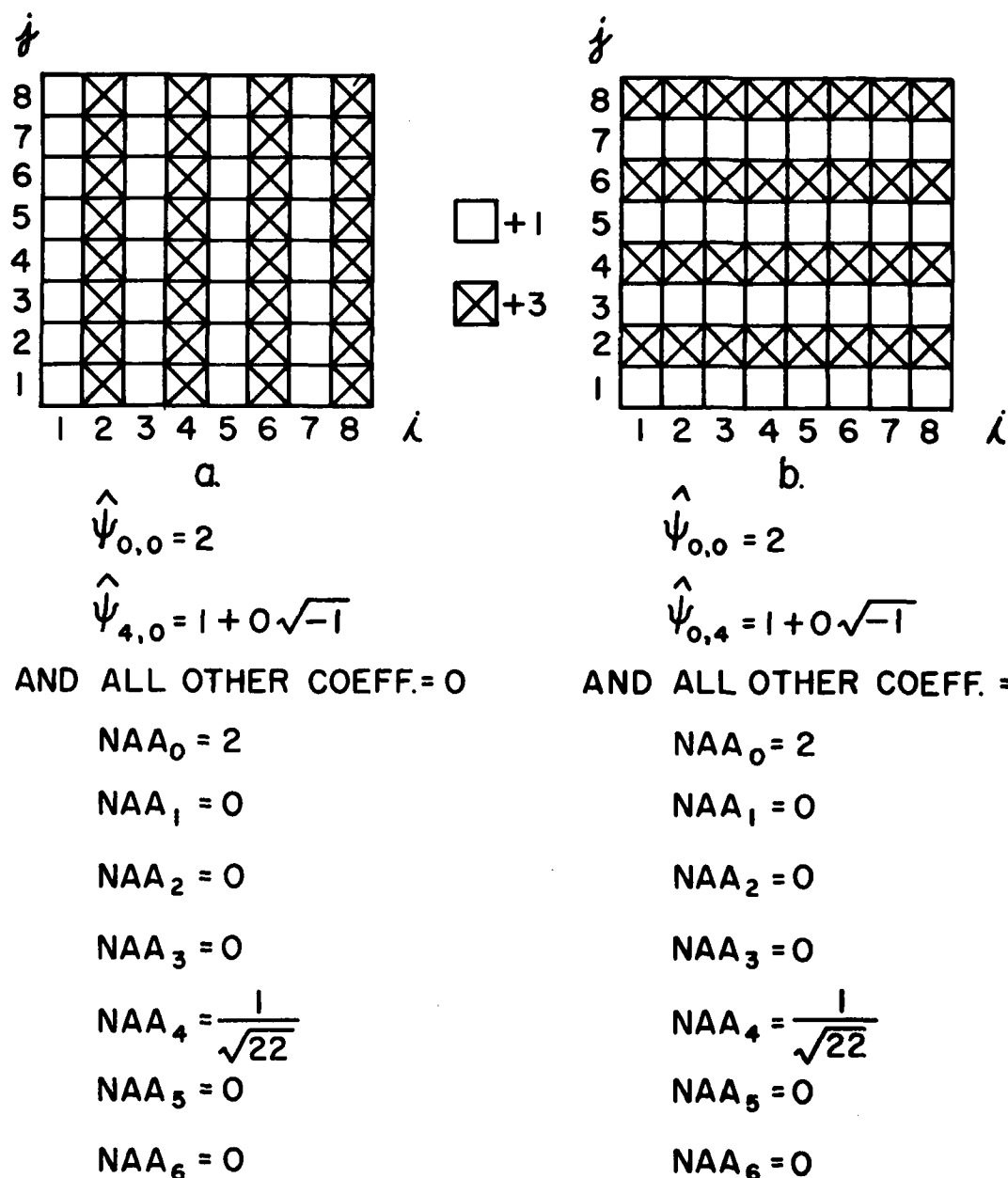
Figure 8. Illustration of the rotation property of the annular integration technique. Array (a) is rotated 90° about (1,1) to obtain (b). Note the pattern in (a) occurs four times horizontally (m = 4) and is unchanged vertically (n = 0); thus $\hat{\psi}_{4,0} = 1$.

Just the opposite is true for (b) (m = 0 and n = 4), hence $\hat{\psi}_{0,4}$ = 1. Although oriented differently each of these two arrays (functions) exhibits the same pattern. The spectral classifier would identify each sample array as the same "type" since their spectra are similar.

3.2 The Discriminant Function

That part of the classifier which determines from what type of cloud a given power spectrum has been generated is called a _Bayes decision rule_. Simply stated, the cloud type chosen by the decision rule to be contained in a satellite imagery sample $\hat{F}$ is the one that is most likely to generate that sample's observed power spectrum. In order to accomplish this, the rule assumes some knowledge of the spectra of every cloud type it is allowed to choose.

The decision rule is allowed to classify a sample as one of nine cloud types: stratocumulus (Sc), stratus (St), cumulus (Cu), clear, altocumulus (Ac), altostratus/nimbostratus (As/Ns)[7], cirrus (Ci), cirrostratus (Cs), and cumulonimbus (Cb). For convenience, these will be referred to as cloud types i = 1,2,3,....,9.

Let the finite set $\Omega = \{\omega_1, \omega_2, \omega_3, \ldots, \omega_9\}$ of the nine _states of nature_ ω_i be the set from which the classifier chooses a cloud type, or in other words, the state of a cloud sample.[8] Since the state of nature is a chance, unpredictable event, ω is called a _random variable_.

The spectral classifier uses the following discriminant function to help the decision rule decide what kind of cloud (which ω_i) a sample contains:

$$d_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)^t \sum_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) - \frac{r}{2} \ln 2\pi - \frac{1}{2} \ln \left| \sum_i \right|$$

$$+ \ln P(\omega_i), \quad i = 1,2,3,\ldots,9, \tag{16}$$

where $\mathbf{x}$ is an r-component column vector[9], $\boldsymbol{\mu}_i$ is the r-component mean vector of class i, t denotes transpose, $\sum_i$ is the r × r class-conditional covariance matrix for class i, $\sum_i^{-1}$ is the inverse of $\sum_i$, $\left| \sum_i \right|$ is the determinant of $\sum_i$,

[7]These two cloud types are merged into one category in order to be consistent with the 3DNEPH.

[8]E.g., ω_3 means the imagery sample contains cumulus clouds.

[9]For this study it is an observed 26-component power spectrum.

and $P(\omega_i)$ is the a priori probability that i occurs.

The matrices $\boldsymbol{\mu}_i$ and $\sum_i$, and the a priori probabilities $P(\omega_i)$, must all be calculated from the sample set of images to be classified prior to any sample classification process which uses (16). The mean power spectrum $\boldsymbol{\mu}_i$ of cloud class i is simply the average of all the power spectra generated from samples containing cloud type i. The matrix $\sum_i$ is composed of the covariances σ_{jk} of NAA_j^i and NAA_k^i, which belong to cloud type i. Along the main diagonal of $\sum_i$ where j = k the covariances σ_{jk} become variances $\sigma_{kk} = \sigma_k^2$ of the kth component NAA_k^i of the spectra that were generated from samples containing cloud type i. $P(\omega_i)$ is simply the relative frequency of occurrence of cloud type i:

$$P(\omega_i) = \frac{N_i}{T} ,$$ (17)

where N_i is the number of samples in the data set to be processed that contain cloud type i, and T is the total number of imagery samples in the data set. Note that $0 \leq P(\omega_i) \leq 1$, and that $\sum_{i=1}^{9} P(\omega_i) = 1$.

Equation (16), except for the term $\ln P(\omega_i)$, is merely the natural log of the multivariate normal density:

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{r/2}\left|\sum\right|^{1/2}} e^{\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^t \sum^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]},$$ (18)

where $\mathbf{x}$, $\sum$, and $\boldsymbol{\mu}$ are as previously defined. Hence d_i is a measure of the probability of a cloud sample's being in state ω_i given that sample's power spectrum $\mathbf{x}$. The i for which d_i is largest is the most reasonable choice of cloud type and one that will be correct most of the time. Such a choice would make $p(\mathbf{x})$ and, therefore, the a posteriori probability $P(\omega_i \mid \mathbf{x})$, the probability of cloud type i given $\mathbf{x}$, largest.

Let $\mathbf{x}$ be a 26-component observed power spectrum generated from a 37 x 37-pixel cloud sample (thus in (16), r = 26). The term $\frac{r}{2} \ln 2\pi = 13 \ln 2\pi$ is merely an additive

constant which is independent of i, and can therefore be ignored. Statistical independence between the NAA's of the power spectra of the same cloud type is assumed. Thus:

$$\sigma_{jk} = \begin{cases} \sigma_k^2 & , \ j = k \\ 0 & , \ j \neq k \end{cases} \qquad \text{for } j,k = 0,1,2,\ldots,25, \quad (19)$$

so that $\sum_i$ reduces to a 26 x 26 diagonal matrix of variances of NAA_k^i, k = 0,1,2,...,25. Rewriting (16) with the above conditions yields:

$$d_i(\mathbf{x}) = -\frac{1}{2} \left[(x_0 - \mu_0^i) \ (x_1 - \mu_1^i) \ldots (x_{25} - \mu_{25}^i) \right] \begin{bmatrix} \frac{1}{(\sigma_0^2)^i} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{(\sigma_{25}^2)^i} \end{bmatrix}$$

$$\begin{bmatrix} x_0 - \mu_0^i \\ x_1 - \mu_1^i \\ \vdots \\ x_{25} - \mu_{25}^i \end{bmatrix} - \frac{1}{2} \ln \left| \sum_i \right| + \ln P(\omega_i)$$

or, simplifying,

$$d_i(\mathbf{x}) = -\frac{1}{2} \sum_{n=0}^{25} \frac{(x_n - \mu_n^i)^2}{(\sigma_n^i)^2} - \frac{1}{2} \ln \left| \sum_i \right| + \ln P(\omega_i),$$

$$i = 1,2,3,\ldots,9, \qquad\qquad\qquad\qquad (20)$$

where $\mathbf{x}$ is an observed power spectrum, a 26-component column vector comprised of x_0, x_1, x_2, ..., x_{25} ($x_n \equiv NAA_n$); μ_n^i is the mean of all NAA_n, class i (from $\boldsymbol{\mu}_i$); $(\sigma_n^i)^2$ is the variance of NAA_n, class i (from $\sum_i$); $P(\omega_i)$ is the a priori probability (17). A d_i is calculated for each of the nine cloud types Sc, St, Cu, ..., Cb. The decision rule decides class k if $d_k > d_i$ for all $i \neq k$.

In the form (20) it is easy to see that a cloud imagery sample's observed power spectrum $\mathbf{x}$ is compared component-by-component to the mean power spectra of each cloud type by measuring the squared **distance** $(x_n - \mu_n^i)^2$ between x_n (NAA_n)

and μ_n^i for all n, all i.[10] The smaller $(x_n - \mu_n^i)^2$ is, the larger d_i is, so that the i for which d_i is largest indicates the cloud type whose mean or "characteristic" power spectrum most closely resembles **x**.

This is how the decision rule uses the discriminant function (20) to make its classification decision. The automated spectral classifier is for this reason called a _minimum distance classifier_. If each $\boldsymbol{\mu}_i$ is thought of as an ideal typical spectrum for cloud class i, then (20) may also be described as a template-matching procedure.

In cases where **x** is "close" to more than one characteristic power spectrum $\boldsymbol{\mu}_i$ (the left-most term on the right side of (20) is or is nearly the same for more than one cloud type i), the classifier's final decision depends on $P(\omega_i)$. For example, when **x** is equally (or nearly equally) close to two mean spectra $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_k$, the choice made by the classifier favors the a priori more likely cloud class. So, if $P(\omega_k) > P(\omega_i)$, then $d_k > d_i$ and cloud type k would be chosen.

A priori probabilities may be used to more suitably adapt the classifier to different regions of the world. For instance, in equatorial regions $P(\omega_9) = P(Cb)$ is much greater than, say, $P(\omega_2) = P(St)$. However, in maritime midlatitude areas, just the opposite is likely the case. Differences in relative frequency of occurrence of cloud types over various geographical areas can be reflected in a proper choice of $P(\omega_i)$'s for each of those areas. This is an advantageous versatility of the spectral classifier technique.

The discriminant function can be thought of as a mathematical formalization of common-sense rules. It is a logical, effective, and straightforward way of determining the cloud type of an imagery sample given that sample's power spectrum.

In addition to using the discriminant function (20) to identify each cloud sample, two other variations of (20) were

---

[10]The numerator of the left-most term on the right side of (20) can be written as $(\mathbf{x} - \boldsymbol{\mu}_i)^t (\mathbf{x} - \boldsymbol{\mu}_i) = \|\mathbf{x} - \boldsymbol{\mu}_i\|$. This is called the _Euclidean distance_ from **x** to $\boldsymbol{\mu}_i$.

utilized to determine how different parts of (20) affect classification decisions.

The mean classifier, which bases its cloud type decisions solely on the mean visible and infrared values (NAA_0) of the cloud samples, simply computes the linear distance between two points:

$$m_i(\overline{VIS}_{obs}, \overline{IR}_{obs}) = \sqrt{(\overline{VIS}_{obs} - \overline{VIS}_i)^2 + (\overline{IR}_{obs} - \overline{IR}_i)^2},$$

$$i = 1,2,3,\ldots,9, \qquad (21)$$

where $\overline{VIS}_{obs}$ and $\overline{IR}_{obs}$ are the observed average visible and infrared values of the sample being classified, and $\overline{VIS}_i$ and $\overline{IR}_i$ are the mean visible and infrared values of all samples of cloud type i. Before (21) can be used, $\overline{VIS}_i$ and $\overline{IR}_i$ must be known for each cloud class i.[11] An m is calculated for each i, and cloud type k is chosen if $m_k < m_i$ for all $i \neq k$. The results obtained by this classification procedure will indicate how much the means of a sample contribute to cloud type classification.

Another variation of (20), which is essentially a curve-fitting operation in a least squares sense, bases its decisions solely on the means and variances of the power spectra of each cloud type:

$$\delta_i(\mathbf{x}) = -\frac{1}{2} \sum_{n=0}^{25} \frac{(x_n - \mu_n^i)^2}{(\sigma_n^i)^2} \quad , i = 1,2,3,\ldots,9, \qquad (22)$$

where $\mathbf{x}$, x_n, μ_n^i, and $(\sigma_n^i)^2$ are as defined for (20). Equation (22) is merely equation (20) without the a priori probability and covariance terms. As in (20), the observed power spectrum $\mathbf{x}$ is compared component-by-component to each mean spectrum $\mathbf{\mu}_i$ by calculating a δ for each of the nine cloud classes. Cloud type k is chosen if $\delta_k > \delta_i$ for all $i \neq k$. When compared with the results of (20), the performance rate of (22) will show how much the classifier relies on a priori probabilities.

[11]The mean classifier follows the same logic as does the spectral classifier.

3.3 Data Availability and Classifier Versatility

During nighttime orbits only infrared cloud data are obtained by satellites. There are also times, albeit less frequent, when only visible data are available for use. Nevertheless, more often than not both visible and infrared data are available for use during the daytime scans of a polar orbiting satellite.

In order to adapt the classifiers (20), (21), and (22) to each of the above conditions, simple and straightforward modifications were made to each.

The visible-only and infrared-only adaptations of the mean classifier (21) are defined as, respectively,

$$m_i^{VIS} = |\overline{VIS}_{obs} - \overline{VIS}_i| \, , \quad i = 1,2,3,\ldots,9, \quad (21a)$$

and

$$m_i^{IR} = |\overline{IR}_{obs} - \overline{IR}_i| \, , \quad i = 1,2,3,\ldots,9, \quad (21b)$$

where $\overline{VIS}_{obs}$, $\overline{VIS}_i$, $\overline{IR}_{obs}$, and $\overline{IR}_i$ are as defined for (21). Formulas (21a) and (21b) make a cloud type decision the same way as (21) does.

The discriminant functions $d_i(\mathbf{x})$ (20) and $\delta_i(\mathbf{x})$ (22) were adapted to separate and simultaneous visible/infrared processing in the same way. The spectral classifiers (20) and (22) use a visible power spectrum $\mathbf{x}$, generated from a visible imagery sample, as input for the visible-only condition, and use an infrared spectrum $\mathbf{y}$ for the IR-only condition. Using (22) as an example, when both visible and infrared imagery are available simultaneously with which to classify a cloud sample, the individual terms $\delta_i(\mathbf{x})$ (visible) and $\delta_i(\mathbf{y})$ (infrared) are calculated separately for each i and then added together. The cloud type i for which the sum $\delta_i(\mathbf{x}) + \delta_i(\mathbf{y})$ is largest is chosen.

4. RESULTS

The available DMSP 5D visible and infrared imagery
was displayed on a McIDAS terminal TV screen. Cloud imagery
samples were then interactively found, examined, and sub-
jectively classified by skilled analysts. Samples contain-
ing mixed clouds were avoided, as were those with coastal
backgrounds and snow-covered areas. Once collected on tape,
a group of cloud samples would then be classified objectively
using the spectral classifier program.

The classifier was allowed to run on two separate sets
of data. The first set, which contained 73 cases, was sub-
jectively classified by a different analyst than was the
second set of 70 samples.

The overall performance of the classifier on both sample
sets was highly similar (see Table 1). This suggests that
the classification technique is sound and consistent. For
these reasons, the results of both trials have been combined
in the following discussion.

TABLE 1. CLASSIFICATION ACCURACY FOR TWO DATA SETS

	Visible	IR	Combined Vis and IR
Mean Classifier, Data Set 1 (73 samples)	34%	37%	49%
Mean Classifier, Data Set 2 (70 samples)	37%	43%	54%
Spectral Classifier, Subset 1	66%	68%	82%
Spectral Classifier Subset 2	64%	63%	79%

The composition of the combined set of cloud samples is
shown in Figure 9. Cirrus is the most common cloud type in
the sample set with 36 of the 143 cases. Cumulus is the next

34
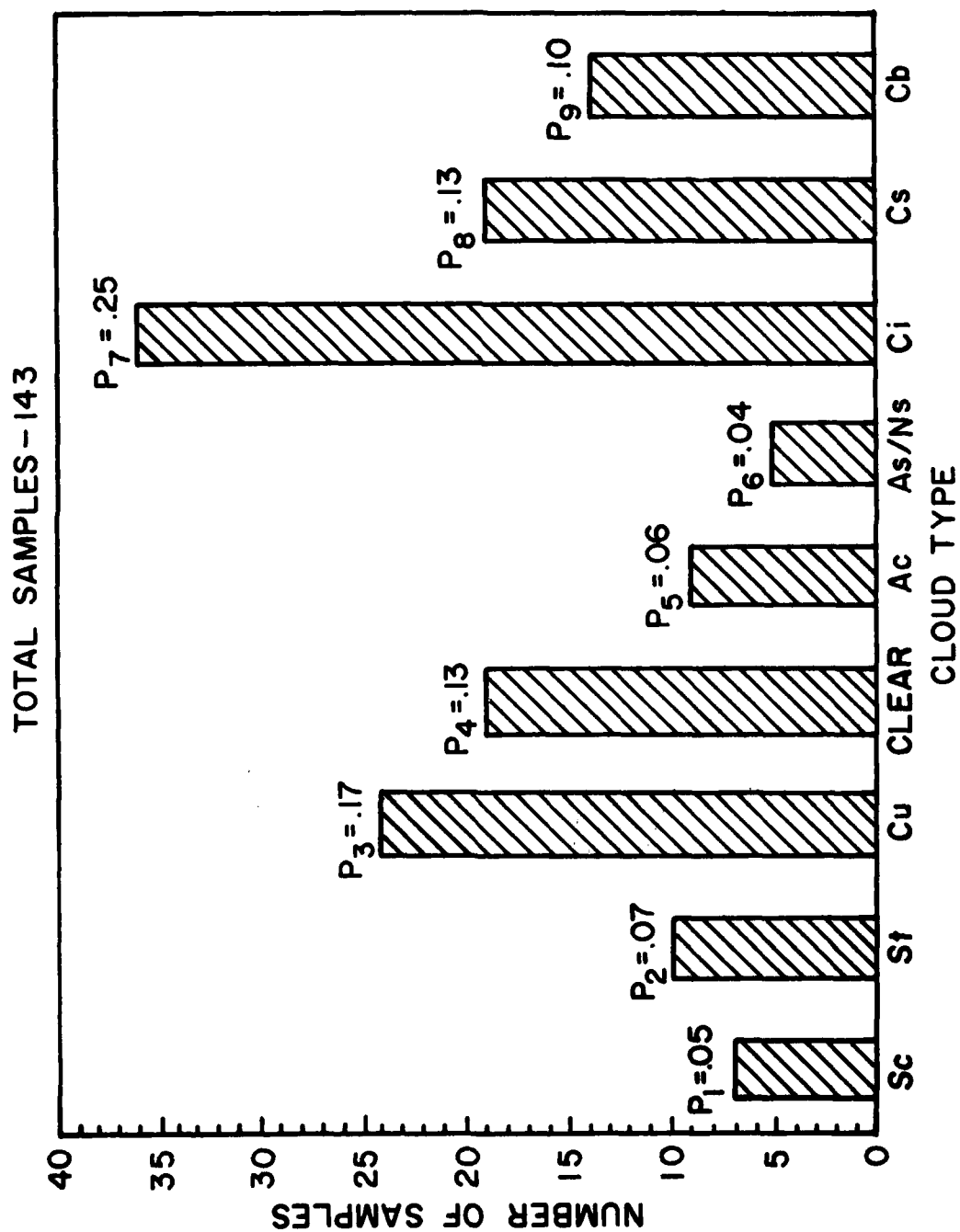
Figure 9. Composition of the cloud imagery sample set.

most frequent with 24 samples. Mid-level clouds (Ac, As/Ns) did not appear too often within the available imagery and thus have a relatively low number of cases: 9 for Ac and 5 for As/Ns. The backgrounds of these cloud samples range from the Mediterranean Sea to the Balkan peninsula, and the deserts of the Middle East and northeast Africa.

With the information shown in Figure 9, the a priori probabilities $P(\omega_i)$ (17) can be computed for each of the nine cloud types. If an observer had to classify this sample set on the basis of this information only, then optimum results would be achieved by choosing the most frequent cloud type in the sample set.

In this case, "blindly" choosing cirrus as the cloud type of all 143 imagery samples would result in a 25% correct classification rate $(P(\omega_7) = P(Ci) = .251)$, since the analyst is sure that 36 of the 143 samples are cirrus. This is considerably higher than a totally random classification approach which yields, on the average, $p = 1/9 = .111$, or an 11% correct classification rate.

The overall results of the three classifiers (20), (21), and (22) should be compared to the above "blind" performance rates as a measure of the classifiers' effectiveness in determining cloud type.

4.1 Mean Classifier

The performance rate of the mean classifier was better than that of the "blind" classification approach (see Table 2a). Attempting to identify from one of the nine cloud types, the mean classifier classified 36% of the cloud samples correctly using visible data only (21a), 40% of the cloud samples using infrared only (21b), and 52% of the cloud samples using visible and infrared data (21). There was a rather large number of misclassifications when this method was utilized, which, although an improvement over the 25% of the "blind observer," implies that the mean of a cloud sample itself is not sufficiently discriminatory to clearly distinguish nine

TABLE 2a. MEAN CLASSIFIER ACCURACY FOR NINE CATEGORIES

Cloud Type	Number of Samples	Visible	IR	Combined Vis and IR
Sc	7	29%	29%	29%
St	10	50%	20%	60%
Cu	24	42%	58%	58%
Clear	19	42%	74%	58%
Ac	9	11%	0	44%
As/Ns	5	60%	40%	60%
Ci	36	25%	17%	42%
Cs	19	47%	42%	63%
Cb	14	29%	64%	50%
Total	143	36%	40%	52%

TABLE 2b. MEAN CLASSIFIER ACCURACY FOR FIVE CATEGORIES

Cloud Group	Number of Samples	Visible	IR	Combined Vis and IR
Low (Sc/St/Cu)	41	54%	66%	71%
Mid (Ac/As/Ns)	14	36%	36%	57%
High (Ci/Cs)	55	44%	25%	53%
Clear	19	42%	74%	58%
Cb	14	29%	64%	50%
Total	143	44%	48%	59%

37

cloud types in a consistent manner. Figure 10 has a plot of visible and infrared averages for each cloud type. Some of the averages are so close (Ac, As/Ns, St) that misclassifications are expected.

A significant factor reflected in these results is the diverse range of backgrounds found among the samples. This greatly affects the means of samples of the same cloud type, especially those in which a lot of background shows through (Cu, clear, Ci). For example, one Cu case has a mean visible count of 19, while another has one of 109, almost six times as high.

When the cloud samples are classified into fewer categories, overall results improve. The nine cloud types were grouped into five: low (Sc, St, Cu), middle (Ac, As/Ns), high (Ci, Cs), clear, and Cb.[12] Thus, for example, a Sc sample classified as Cu was considered a correct identification, or "hit."

Using visible means only, low clouds were correctly hit 54% of the time, middle clouds only 36% of the time, and high clouds 44% of the time. For all five categories, 44% were classified correctly.

Using infrared only, 48% hits were achieved: 66% low clouds were hit, 36% middle clouds were hit, but (surprisingly) only 25% high clouds were correctly classified.

When visible and infrared means were used simultaneously, 71%, 57%, and 53% of the low, middle, and high cloud samples, respectively, were correctly classified for a combined five-category 59% hits.

If the above five-category results, listed in Table 2b, are compared with the "blind" classification approach performance rate of 38% (attainable by calling each sample

---

[12]Cs and Cb were confused at times by the classifiers since they often look very similar in the visible and in the infrared. Because of the severe weather frequently associated with Cb, it was felt that these two cloud types should be in separate classes even though results would improve if they were grouped together.
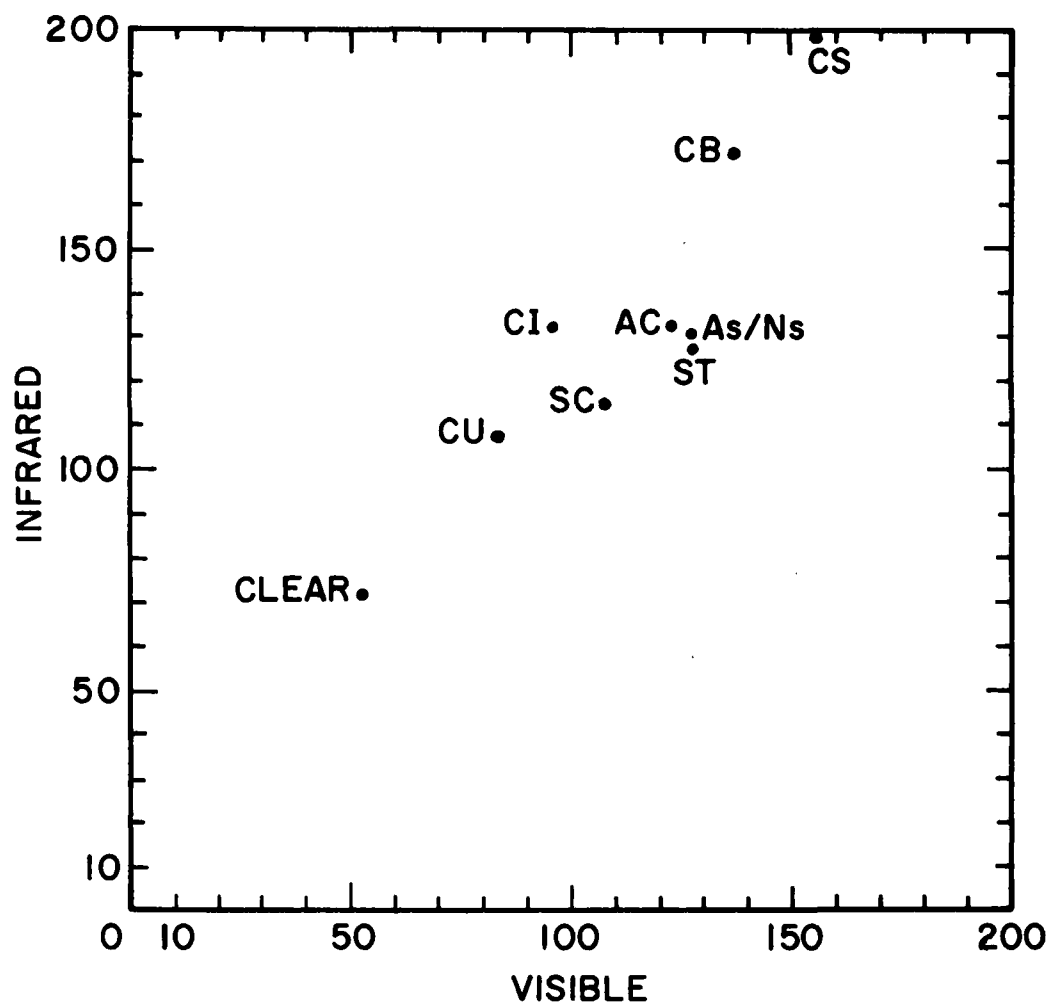
Figure 10. Average visible and infrared means for each cloud type, sample subset 2 (70 samples). High values on the infrared scale represent cold temperatures.

"high"), then it is seen that the mean classification approach
has an advantage over chance.

4.2 Average Power Spectra Characteristics

This section discusses the results obtained when the
cloud type decision rules (20) and (22) were utilized on the
samples. Before these two methods could be applied, the means
and variances of the power spectra of each cloud type were
computed.

The average visible and infrared power spectra of each
of the nine cloud types contained in sample set 2 (70 cases)
are plotted in Figures 11 and 12. As can be seen in these
graphs, the most noticeable differences among the average
spectra lie in the lower wavenumbers, and especially in the
means (NAA_0 (15)) of the cloud types. Cs has the highest
visible and infrared means, while Cb has the second highest.
Appropriately, clear has both the lowest visible and lowest
infrared means. (High values on the infrared scale corres-
pond to low temperatures while low values correspond to high
temperatures.)

As previously stated, coefficients of a Fourier expan-
sion are the amplitudes of waves which lie about the mean of
the function they describe (see Figure 3). They are therefore
some measure of how much the values of that function deviate
from its mean. Hence coefficients $\hat{\psi}_{m,n}$ that lie in the
smaller-radius (lower ρ) annular bands are the amplitudes
of larger-scale variations (larger λ - see equation (14))
in cloud brightness and spatial configuration, and the coeffi-
cients that lie in the higher ρ annular bands are the ampli-
tudes of smaller-scale variations (smaller λ). It is in
this sense that the average power spectra, generated from
Fourier coefficients, are indicative of the amount of texture
present in the imagery field. As the higher wavenumber NAA's
increase, the imagery sample appears less uniform.

Beyond wavenumber 6, the range of values (the differ-
ence NAA_{max} - NAA_{min}) for the average spectra of the strati-

40
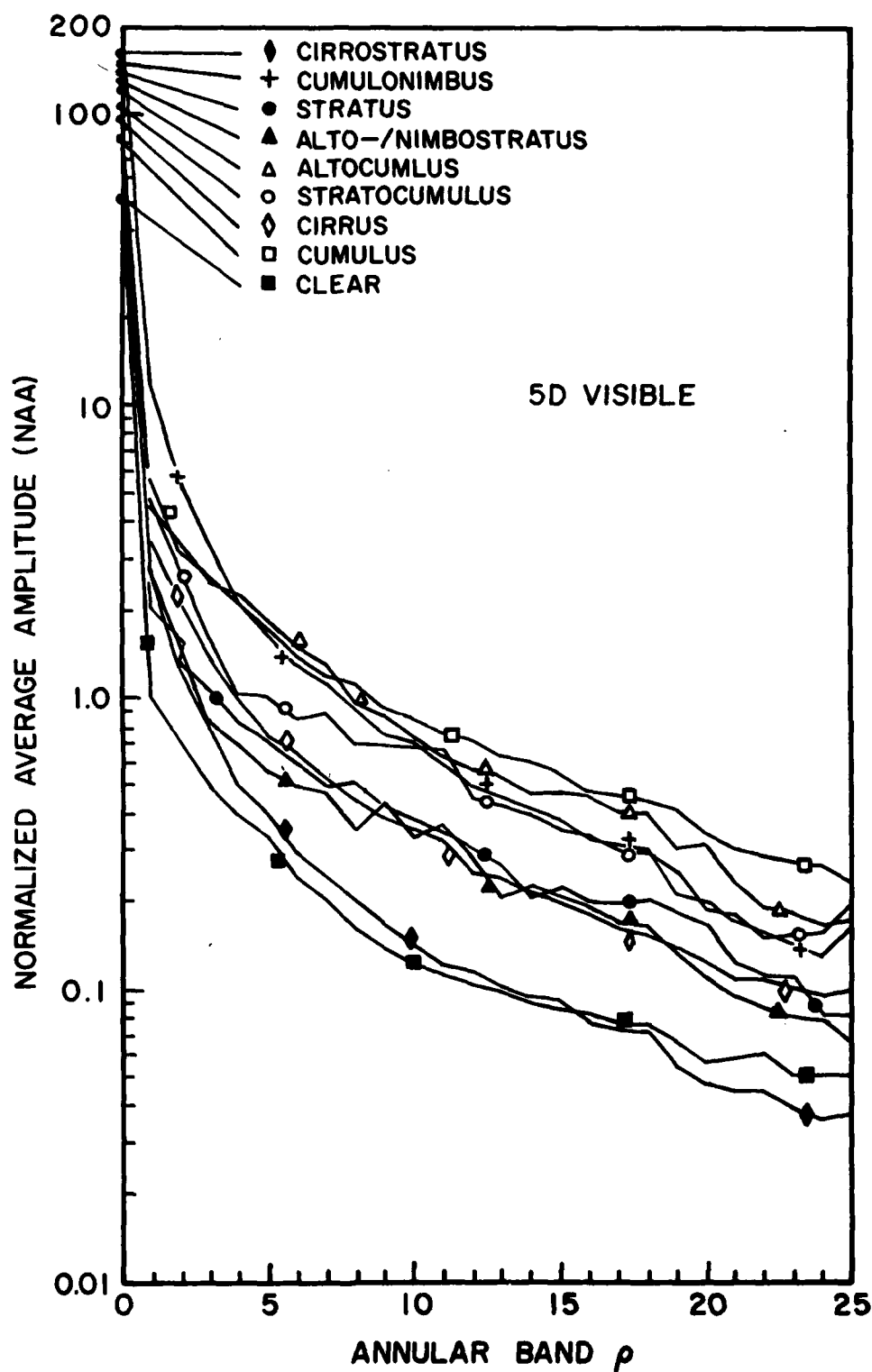
Figure 11. Semi-log plot of visible average power spectra of each cloud type, sample subset 2.

41
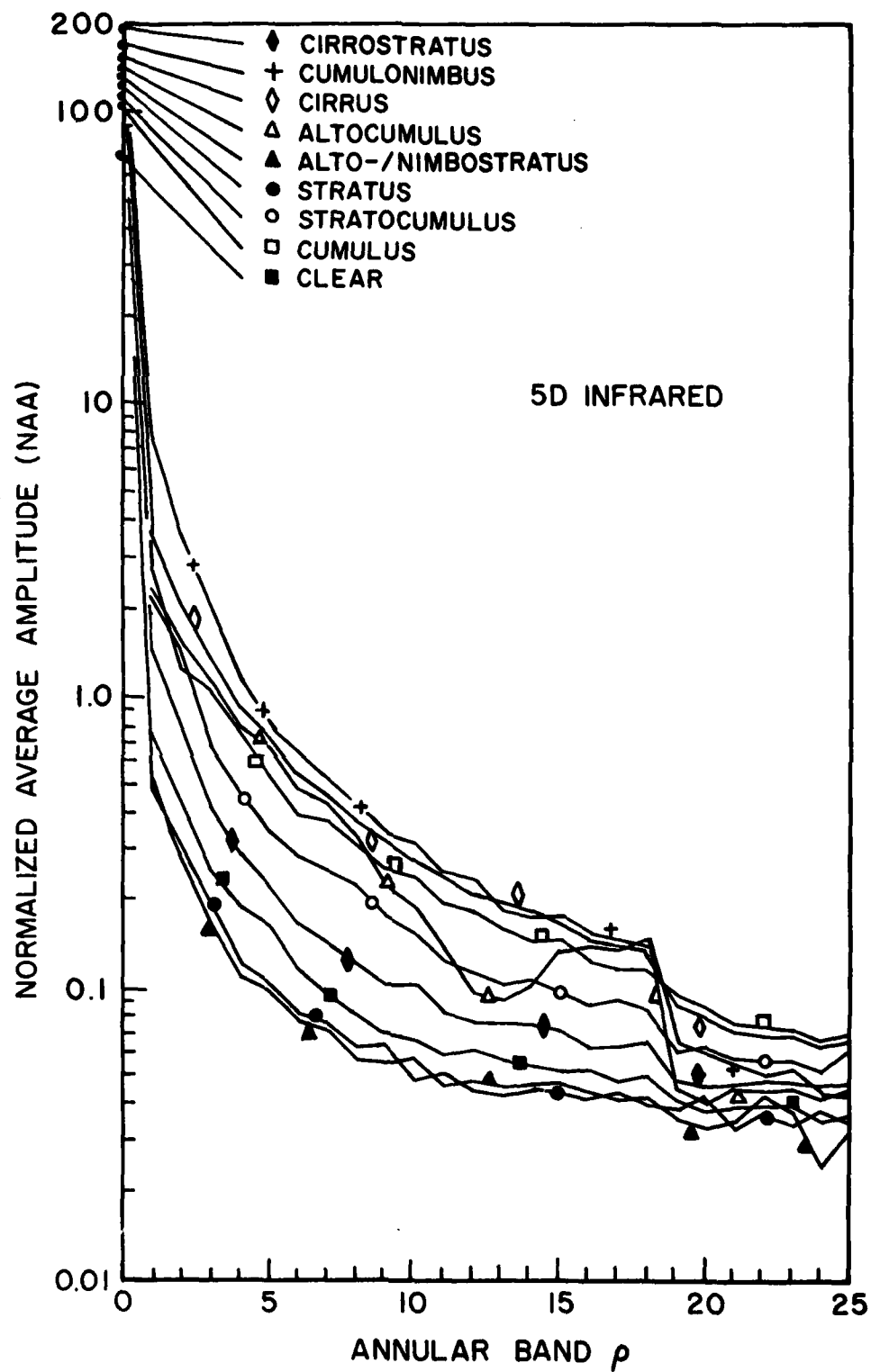
Figure 12. Semi-log plot of infrared average power spectra of each cloud type, sample subset 2.

42

form cloud types (As/Ns, St, Cs) is particularly small, especially in the infrared, indicating that their imagery is quite smooth in appearance. On the other hand, the corresponding range of the cumuliform cloud type spectra (Cu, Sc, Ac) is significantly larger, especially in the visible.[13] This implies that cumuliform imagery is much less uniform in appearance than is stratiform imagery. Figures 13 and 14 are plots of average spectra for wavenumbers greater than 6.

4.3 The Discriminant Functions

The Bayesian decision rules (22) and (20), as expected, performed much better than the mean classifier. Improvements of 20 to 30% over the mean classifier performance rates were obtained when the discriminant functions were used. Such a significant increase in classifier accuracy is mostly due to the fact that there is a lot more cloud feature information present in the power spectrum of a cloud sample than there is in just its mean. Results are shown in Tables 3 and 4.

The largest increase observed was one of 32%. Using visible data only, the accuracy of the mean classifier was 36%, whereas that of the minimum distance classifier (20) was 68%.

In general, Cs was the cloud type most often correctly classified, especially by (20). When both visible and infrared data were used to determine cloud type, 17 of 19 Cs cases were properly identified (89%). Cb cloud samples were also well classified by both (20) and (22).

Mid-level cloud types were the most inconsistently classified. For example, only two Ac samples were called Ac by (22) when visible was used alone. However, when visible

[13]Stratiform range $\rho > 6$; Vis: $\approx .40$, IR: $\approx .04$.
 Cumuliform range $\rho > 6$; Vis: ≈ 1.0, IR: $\approx .40$
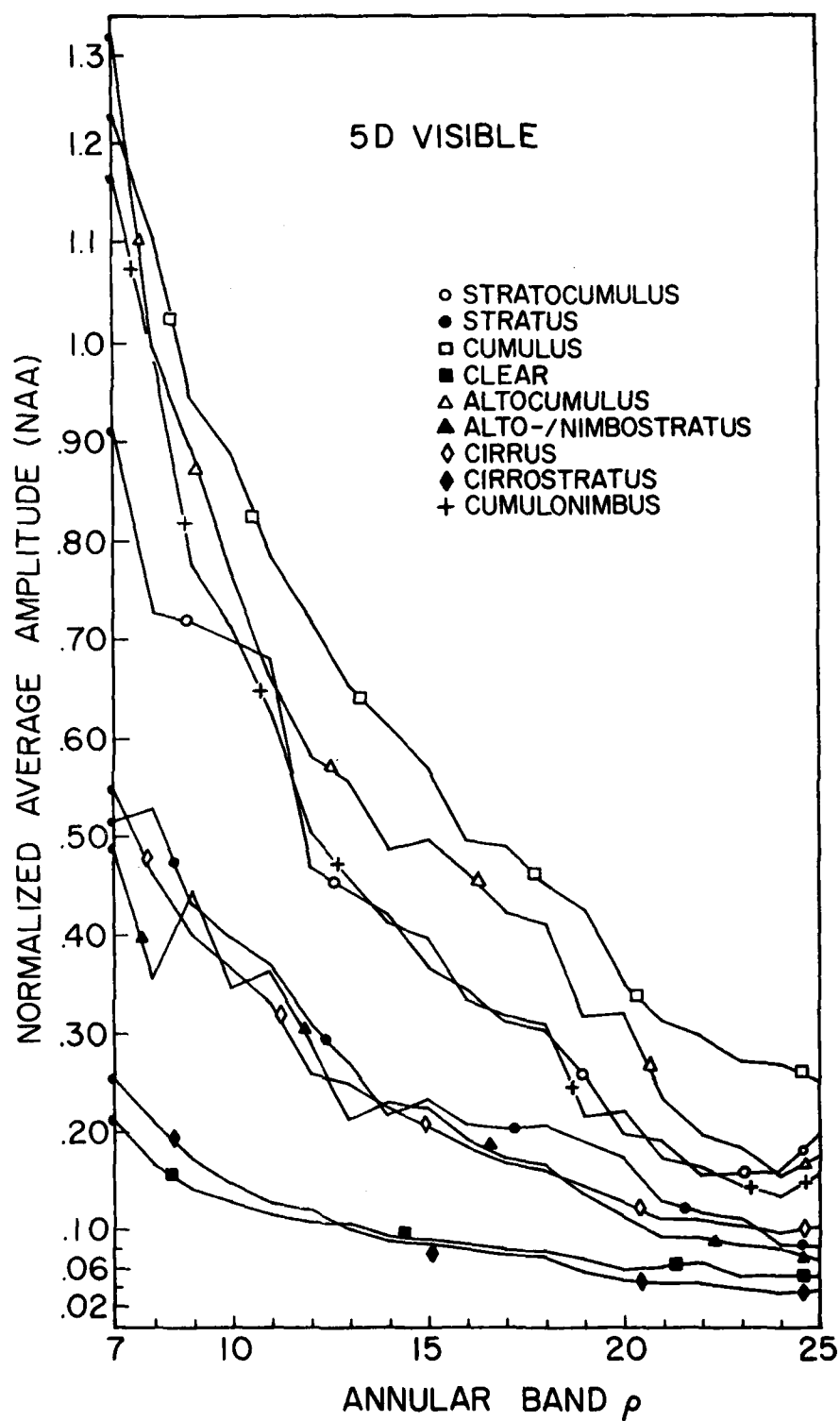
Figure 13. Linear plot of visible average power spectra of each cloud type beyond annular band 6, sample subset 2.
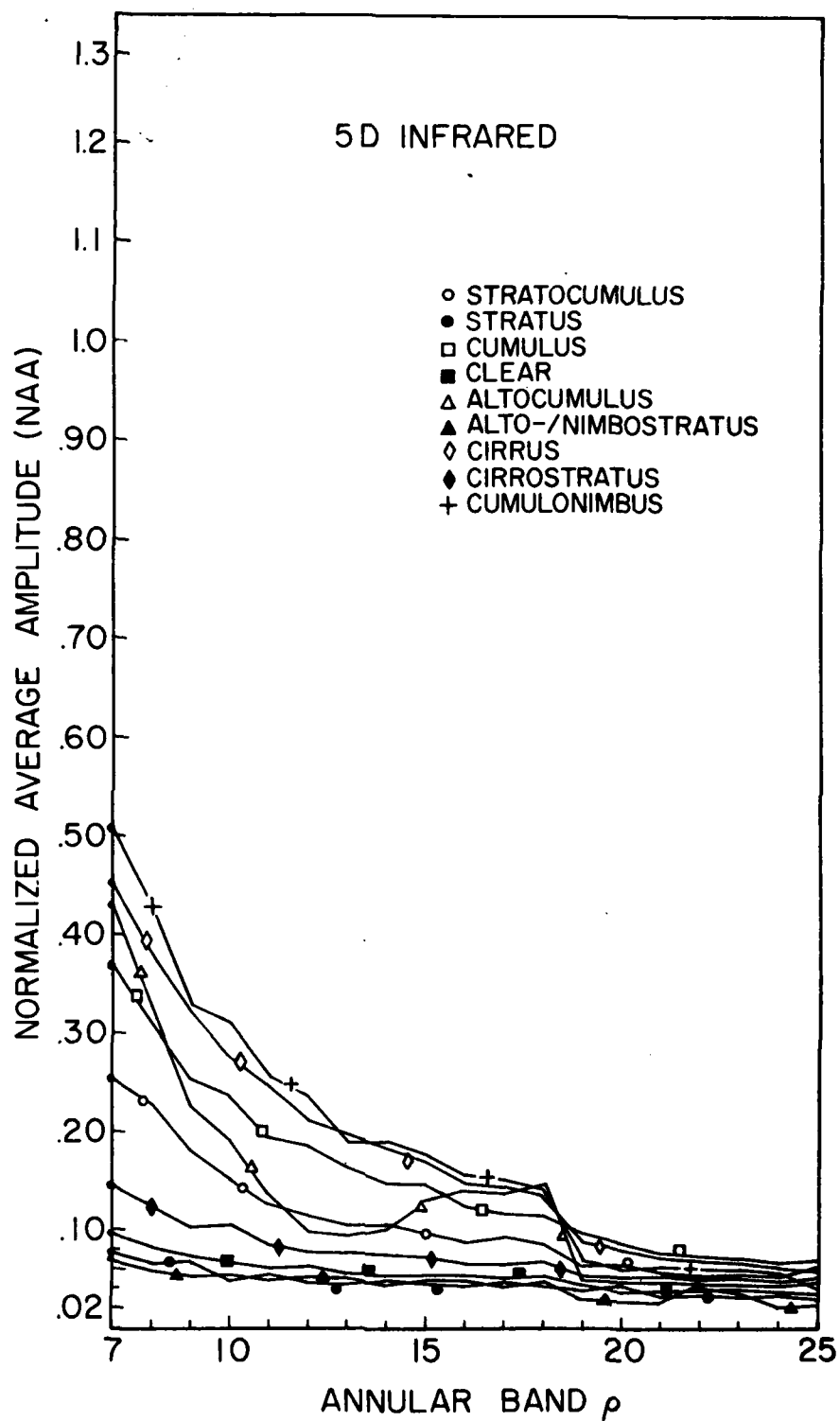
Figure 14. Linear plot of infrared average power spectra of each cloud type beyond annular band 6, sample subset 2.

TABLE 3. NINE-CATEGORY SPECTRAL CLASSIFIER ACCURACY WITHOUT
A PRIORI AND COVARIANCE TERMS (22)

Cloud Type	Number of Samples	Visible	IR	Combined Vis and IR
Sc	7	57%	29%	86%
St	10	40%	20%	70%
Cu	24	67%	54%	79%
Clear	19	74%	74%	74%
Ac	9	22%	33%	67%
As/Ns	5	20%	0	100%
Ci	36	42%	72%	83%
Cs	19	74%	84%	89%
Cb	14	79%	79%	86%
Total	143	57%	61%	81%

TABLE 4. NINE-CATEGORY SPECTRAL CLASSIFIER ACCURACY WITH
A PRIORI AND COVARIANCE TERMS (20)

Cloud Type	Number of Samples	Visible	IR	Combined Vis and IR
Sc	7	43%	43%	71%
St	10	70%	80%	100%
Cu	24	58%	50%	71%
Clear	19	95%	89%	89%
Ac	9	78%	44%	78%
As/Ns	5	100%	100%	100%
Ci	36	44%	72%	64%
Cs	19	84%	74%	89%
Cb	14	79%	71%	86%
Total	143	68%	69%	79%

alone was used by (20), 7 of 9 Ac cases were correctly classified. A similar statement can be made of As/Ns samples using their infrared spectra only (0 out of 5 vs. 5 out of 5). Of course, performance of the classifier on Ac and As/Ns cloud types should not be judged as either poor or good on the basis of this information alone, since only a small number of these mid-level samples were selected and processed.

On the other hand, clear samples were most consistently classified by the discriminant functions, as can be seen in Table 3 (74%) and Table 4 (89%).

The mistakes the classifier made when it erred were often reasonable. Oftentimes Sc was mistaken for Cu, or Cs for Cb. A measure of how often such mistakes occur lies in comparing the discriminant function nine-category results with the discriminant function five-category (low, middle, high, clear, Cb) results.

The overall five-category combined performance rates increased up to 10% over those of the nine-category set. This implies that a substantial number of samples misclassified as one of nine types were not misclassified as one of five types. Results are shown in Tables 5 and 6.

The improvement of the accuracy of the discriminant functions when allowed to classify from five types instead of nine should not be misinterpreted. Five-category accuracies cannot get worse. By retabulating the results for nine cloud categories into five, the overall results of the classifier can only remain the same or get better. Therefore, the important point is not whether the classifier has improved, but by how much it did improve. Such a comparison indicates the amount of "reasonable" errors the classifier makes.

The performance of the visible-only nine-category classifier (22), poorest of all at 57% (see Table 3), improved by nearly 10% to 66%. This was better than any of the other observed increases.

The classifier (20) called a sample from the low cloud category either Sc, St, or Cu 80% of the time, using infra-

TABLE 5. FIVE-CATEGORY SPECTRAL CLASSIFIER ACCURACY WITHOUT
A PRIORI AND COVARIANCE TERMS (22)

Cloud Group	Number of Samples	Visible	IR	Combined Vis and IR
Low (Sc/St/Cu)	41	71%	54%	83%
Mid (Ac/As/Ns)	14	21%	21%	79%
High (Ci/Cs)	55	69%	84%	89%
Clear	19	74%	74%	74%
Cb	14	79%	79%	86%
Total	143	66%	67%	84%

TABLE 6. FIVE-CATEGORY SPECTRAL CLASSIFIER ACCURACY WITH
A PRIORI AND COVARIANCE TERMS (20)

Cloud Group	Number of Samples	Visible	IR	Combined Vis and IR
Low (Sc/St/Cu)	41	66%	80%	85%
Mid (Ac/As/Ns)	14	86%	64%	86%
High (Ci/Cs)	55	75%	75%	85%
Clear	19	95%	89%	89%
Cb	14	79%	71%	86%
Total	143	76%	77%	86%

red spectra. It also correctly classified three of 7 Sc samples, 8 of 10 St samples, and 12 of 24 Cu samples from nine cloud types using infrared spectra. Of the 18 samples (4 Sc, 2 St, and 12 Cu) that were not correctly classified from nine types, 10 of them were placed within the low cloud group. Thus more than half of the previous incorrect classifications were recovered because of five-category partitioning.

Although high clouds had the highest hit percentage, in general both high and low clouds had similar performance rates. As was the case with the nine cloud types, mid-level clouds were inconsistently classified.

4.4 Significance of A Priori Knowledge and Its Effect on the Classifier

Equation (20) differs from (22) by the "a priori" terms

$-\frac{1}{2} \ln \left| \sum_i \right|$ and $\ln P(\omega_i)$. These extra terms are intended to

discriminate between two or more cloud types whose mean spectra μ_i are all "close" to an observed sample spectrum x. The difference in the results of using (20) and (22) was most significant when using visible data only. Increases of 11 and 10% were observed for the nine category (57 to 68%) and five category (66 to 76%) sets. All other percentages increased on the order of 2-8%.

The degree of help that the a priori terms give the classifier should not be taken for granted since it may be unique to this data sample set. More than half of the two largest increases of 11 and 10% were due to a 64% increase in the number of mid-level clouds correctly classified when (20) was used. As described earlier, mid-level clouds were the most inconsistently classified group, so that such a marked increase should be carefully considered before determining the significance of these results.

In those cases when the classification rate of the mid-level clouds did not change so abruptly (IR-only, combined visible and IR), the more modest improvements of 2-8% do not

overwhelmingly suggest the use of (20) over (22). It is true
that the results obtained by (20) were almost always better,
but the processing of a larger data set would provide better
information with which to determine the most favorable dis-
criminant function. The a priori probability and covariance
terms in (20) for such a data set would tend to be a better
climatology of cloud types.

4.5 The Classifier vs. Chance

The spectral classifier is clearly advantageous over a
chance classification approach. Choosing the highest a pri-
ori probability from nine cloud types yields a 25% correct
classification rate while the automatic classifier yields an
81% correct classification rate. Similarly, choosing the
highest a priori probability from five cloud types yields a
38% correct classification rate while the automated classifier
yields an 86% correct classification rate.

In a previous similar study, Blackman and Pickett (1979)
obtained a 46% classification rate for a six-category set
consisting of the Sc, Ac, Ci, Cs, and Cb cloud types. Only
DMSP visible data of 1/3 nm resolution (along the satellite
subtrack[14]) were available for this study, and only the dis-
criminant function (20) was used. Considering only these six
cloud types, this study's corresponding correct classification
rate (using 2/3 nm visible data only, along with (20)) was
62%.

4.6 Another Data Reduction Technique

The question of how little information is actually
necessary to enter into a decision-making process in order
to achieve desirable results is highly important since prac-
tical considerations require that the great volume of satel-

[14]Refer to Bunting and Fournier (1980) for a more complete
description of DMSP satellite sensor capabilities.

lite data be substantially reduced. This technique has
helped that problem by reducing the number of data values
from 1,369 (an input imagery sample) to 26 (its power spec-
trum). However, there are variations of this method that can
further reduce the amount of information it needs to effec-
tively classify cloud samples.

It would be unwise to discard any parts of the original
imagery array itself, since the transform of a space-depen-
dent cloud image with "holes" in it[15] could not be computed.
But once the imagery is transformed into its Fourier coeffi-
cients, the following two possibilities are left open: (1)
utilize less than 26 components of a power spectrum in the
final decision process (discriminant function); and/or (2)
decrease the number of coefficients used in computing a
cloud sample's power spectrum.

The first option requires a more detailed study of the
power spectra generated from the transforms of the imagery
samples of each cloud type. Blackman and Pickett (1979)
have suggested that most of the class-discriminating power
lies in only a few of the 26 spectral components, so that not
all of them are necessary to obtain the results already pre-
sented. This possibility was not studied but it deserves
added attention.

The second option was studied by arbitrarily decreas-
ing the number of coefficients in computing a power spectrum.
Only those coefficients $\hat{\psi}_{m,n}$ with non-negative subscripts
$m,n \geq 0$ were allowed into any power spectra calculations.
All others were ignored. (A more complete description is
given in Appendix C.) This caused minimal discriminant
function modifications, since a full 26-component spectrum
could still be generated from these fewer coefficients. By
restricting the coefficients used to those which lie in the
first quadrant of the principal part of the wavenumber plane
($m,n \geq 0$; see Figure 6), at least one Fourier coefficient
lies in any given annular band generated by (10). This im-
plies that at least some information about every harmonic

[15](i,j) points where the input function $\hat{F}(i,j)$ would be undefined.

51

(wavenumber) present in the cloud imagery field will be included in its spectrum.

This reduction method cut the number of Fourier coefficients processed for the power spectrum of a 37 x 37 array from 1,369 to 361, a drop of almost 4:1. The advantages here are computational speed and storage reduction, but the disadvantage may be loss of information. The savings in computational time are small, since annular integration is much quicker than computing the FFT. Thus a slight to no decrease in the performance rate of this technique with respect to the "all-coefficient" method would justify its use.

Results for the first quadrant coefficient method are shown for both the nine and the five category trials in Tables 7, 8, 9, and 10. These Tables are cross-referenced to their corresponding "all-coefficient" method results. The number in parentheses which immediately follows each trial's overall nine- or five-category performance rate (e.g., 60% (+3%), Vis, Table 7) denotes the change in classifier accuracy between the first quadrant or "quarter-coefficient" method and its corresponding "all-coefficient" method. For example, 60% (+3%) indicates that the quarter-coefficient method correctly classified 3% more samples than did the corresponding all-coefficient classifier.

When using (22) to classify the cloud samples, the quarter-coefficient method achieved slightly different results than did the all-coefficient method. There were no significant changes in any performance rates (±2 or 3 samples), suggesting that it is just as well to use (22) with power spectra generated from either all or just one-fourth the coefficients in the Fourier transforms of each imagery sample.

When using (20) and only visible data to classify the cloud samples, the quarter-coefficient method achieved higher results (+7 to +12 samples) than did the all-coefficient method. Combined visible and infrared classifier performance also increased, but at a more modest rate (+4 samples).

The results for this data set favor the use of (20) with the power spectra generated from just one-fourth the coeffi-

TABLE 7.   NINE-CATEGORY SPECTRAL CLASSIFIER ACCURACY WITHOUT
A PRIORI AND COVARIANCE TERMS (22)

"QUARTER-COEFFICIENT" TRIAL
(Compare with Table 3)

Cloud Type	Number of Samples	Visible	IR	Combined Vis and IR
Sc	7	57%	29%	71%
St	10	40%	20%	80%
Cu	24	63%	71%	79%
Clear	19	74%	79%	79%
Ac	9	33%	56%	56%
As/Ns	5	0	0	80%
Ci	36	56%	72%	83%
Cs	19	74%	79%	89%
Cb	14	86%	71%	86%
Total	143	60%(+3%)	64%(+3%)	80%(-1%)

TABLE 8.   NINE-CATEGORY SPECTRAL CLASSIFIER ACCURACY WITH
A PRIORI AND COVARIANCE TERMS (20)

"QUARTER-COEFFICIENT" TRIAL
(Compare with Table 4)

Cloud Type	Number of Samples	Visible	IR	Combined Vis and IR
Sc	7	43%	43%	71%
St	10	70%	70%	100%
Cu	24	75%	46%	79%
Clear	19	95%	95%	89%
Ac	9	78%	22%	67%
As/Ns	5	100%	100%	100%
Ci	36	58%	69%	72%
Cs	19	89%	79%	89%
Cb	14	86%	71%	86%
Total	143	76%(+8%)	67%(-2%)	82%(+3%)

TABLE 9. FIVE-CATEGORY SPECTRAL CLASSIFIER ACCURACY WITHOUT
A PRIORI AND COVARIANCE TERMS (22)

"QUARTER-COEFFICIENT" TRIAL

(Compare with Table 5)

Cloud Group	Number of Samples	Visible	IR	Combined Vis and IR
Low (Sc/St/Cu)	41	66%	59%	83%
Mid (Ac/As/Ns)	14	21%	36%	64%
High (Ci/Cs)	55	73%	80%	89%
Clear	19	74%	79%	79%
Cb	14	86%	71%	86%
Total	143	67%(+1%)	69%(+2%)	83%(-1%)

TABLE 10. FIVE-CATEGORY SPECTRAL CLASSIFIER ACCURACY WITH
A PRIORI AND COVARIANCE TERMS (20)

"QUARTER-COEFFICIENT" TRIAL

(Compare with Table 6)

Cloud Group	Number of Samples	Visible	IR	Combined Vis and IR
Low (Sc/St/Cu)	41	71%	71%	90%
Mid (Ac/As/Ns)	14	86%	50%	79%
High (Ci/Cs)	55	82%	78%	91%
Clear	19	95%	95%	89%
Cb	14	86%	71%	86%
Total	143	81%(+5%)	75%(-2%)	89%(+3%)

cients of the Fourier transforms of each imagery sample.
However, the use of some or all available Fourier coeffi-
cients should not be confidently determined on the basis of
these results alone since the data sample is not large. As
previously mentioned, time savings are small. Moreover, no
more than 7% of the samples were reclassified as something
else when one-fourth of the coefficients were used.

4.7 The Spectral Classifier and 3DNEPH

The discrete Fourier transform spectral analysis al-
gorithm was developed as a potential alternative cloud classi-
fication technique to that part of 3DNEPH which determines
cloud types. Table 11 contains the spectral classifier (20)
performance rates (all-coefficient method) for those cloud
categories used by 3DNEPH. Only Cb clouds are classified by
3DNEPH when using visible satellite data. Four cloud groups,
Cs/Cb, As/Ns, Ci, and Ac, are used when only infrared data
are available to 3DNEPH. (A slash (/) separating two cloud
types within a group indicates that those clouds are indistin-
guishable to 3DNEPH.) The eight cloud types that the DFT
classifier uses are also used by 3DNEPH when both visible and
infrared data are utilized, with the clear case omitted.

On the average, the DFT spectral classifier (20) was
able to correctly identify 7-8 cloud samples out of every
10 it was allowed to process.

5. CONCLUSIONS

The automated DFT spectral classifier was allowed to
classify 143 cloud imagery samples. The results obtained
from such a small data set left some questions unanswered.
Nevertheless, they are encouraging.

The classifier's performance rate was similar for two
cloud sample subsets, suggesting that it is consistent in its
imagery classification decisions and that the classifier
yields reproducible results.

TABLE 11.   CLASSIFIER ACCURACY FOR 3DNEPH CLOUD CATEGORIES

Cloud Type	Number of Samples	Number of Correct Identifications	% Correct
VISIBLE			
Cb	14	11	79%
INFRARED			
Cs/Cb	33	26	79%
Ci	36	26	72%
As/Ns	5	5	100%
Ac	9	4	44%
Total	83	61	73%
COMBINED VISIBLE AND INFRARED			
Sc	7	5	71%
St	10	10	100%
Cu	24	17	71%
Ac	9	7	78%
As/Ns	5	5	100%
Ci	36	23	64%
Cs	19	17	89%
Cb	14	12	86%
Total	124	96	77%

By partitioning individual cloud types into larger categories of related types, many of the classification errors were shown to be classifications to related cloud types (e.g., Sc for Cu).

The use of infrared spectra to determine a sample's cloud type proved helpful. When used alone, the IR-only classifier correctly classified 65-70% of the imagery samples in the data set. This is an important result, since the use of only IR data is forced over regions underneath either a nighttime satellite pass or a low sun angle daytime pass. As hoped, the use of infrared data in conjunction with visible data greatly enhanced the reliability of the classifier.

The following are offered as suggestions for potential improvements to the automated cloud imagery classification algorithm:

1. The averaged 2/3 nm DMSP data used by this routine provided a reasonable spatial resolution. The unaveraged 1/3 nm data undoubtedly contains enhanced small-scale cloud feature distinctions and should be tested; however, its use may not yield significantly better results. The use of 1/3 nm imagery would require the routine to take in 75 x 75-pixel arrays (25 x 25 nm earth coverage). This represents a jump from $37 \cdot 37 = 1,369$ to $75 \cdot 75 = 5,625$ input values, an increase of over four times as much data that would have to be processed by the routine. The cloud identification accuracy of the classifier could not increase fourfold with the implementation of 1/3 nm data, while data processing time would, since such an increase in the amount of input would place additional burden on computer time and storage allocations.

2. The collection of much larger imagery sample sets, adequately representative of all cloud types, would firmly test the classifier's ability to distinguish among cloud types.

3. The many background features that show through the clouds in imagery samples, especially those containing high,

57

thin cirrus or fair weather cumulus clouds, cause samples with the same cloud type to look different to the classifier if their backgrounds are different. One way to minimize this adverse effect would be to replace all cloudless pixels with a brightness value of 0 (totally black pixel). Samples would then have a common background, the difference among all samples being only the cloud types in them. This would assure that only clouds will differentiate one sample's cloud type from another's.

4. The spectral classifier had a performance rate nearly twice that of the mean classifier, but the spectral classifier's discriminant function required 26 times more information. It is most likely that not all 26 components of the power spectrum of a 37 x 37 pixel cloud imagery sample need be used by a discriminant function to obtain similar accuracy results. Perhaps adequate discriminating power among the cloud types lies in using fewer spectral components, the only question being which components to use and which not to use.

5. The main feature used by this classification algorithm is a power spectrum. An imagery sample's power spectrum is assumed to contain enough information about any clouds within the sample to identify the sample's cloud type. This was obviously not always true.

It was observed that beyond the sixth or seventh band number ρ, the characteristic spectra of all cloud types had dissimilar, distinct ranges of values (NAA_{max}^i - NAA_{min}^i). When compared with these ranges, a sample's corresponding observed spectral range can provide additional useful information to the automated decision process, information which can successfully identify a sample's cloud type in cases where that sample's spectrum is not sufficient to correctly classify its cloud type.

6. As wavenumbers increase, NAA's tend to decrease in magnitude. Consequently, the variances of the higher annular band number ρ NAA's also tend to be small. In this study, generally over 90% of the variances of the higher bandnumber components of power spectra were significantly

58

less than 1.

Therefore, the product $\left|\sum_i\right|$ (determinant of $\sum_i$) of all these variances (see equation (20)) was an extremely small positive number, in turn making the term $-\frac{1}{2} \ln\left|\sum_i\right|$ in (20) rather large. It was often greater than any other term in (20), including $\ln P(\omega_i)$. Such dominance does not seem to give $\ln P(\omega_i)$ the weight it should have in the decision-making process of the classifier. For this reason, $-\frac{1}{2} \ln\left|\sum_i\right|$ is better left out of (20) or scaled down to a more reasonable magnitude. The a priori probabilities $P(\omega_i)$, since $\in [0,1]$, kept the term $\ln P(\omega_i)$ reasonable in magnitude and should remain untouched as a part of the discriminant function.

7. The use of independent data, i.e., data which have not been used to calculate μ_i and $\sum_i$ (see (16)), would determine how sound the logic of the classifier is. Mean spectra and covariance matrices can be generated from an imagery sample set well representative of all cloud types in frequency of occurrence and in texture. These spectra could then be used with (20) to objectively classify different sets of imagery samples, and would eliminate the need for their subjective classification (unless desired for independent verification purposes). In its current state the classifier computes μ_i and $\sum_i$ directly from the input data set, in addition to requiring interactively supplied input in the form of subjective imagery sample cloud type identification.

8. In this study, the visible and infrared spectra of a cloud sample were used simultaneously, but independently, to identify the sample's cloud type. The use of cross-correlation and covariance information between the visible and infrared spectra of an imagery sample, as first suggested by Blackman and Pickett (1979), may prove to be a sounder method of using both data types. Refer to Appendix E for a description of this method.

On the basis of this study's results, the spectral classifier has established itself as a fast, effective, and consistent satellite imagery classification technique. But as with any prediction technique, this one will ultimately

succeed or fail depending on whether or not the results it obtains are found through future experiments to be both accurate and reliable.

REFERENCES

E. S. Blackman and R. M. Pickett, "Automated Processing of Satellite Imagery Data: Test of a Spectral Classifier," AFGL-TR-79-0040, 1979, 33 pp.

E. O. Brigham, The Fast Fourier Transform, Prentice-Hall Inc., Englewood Cliffs, 1974, 252 pp.

J. T. Bunting and R. F. Fournier, "Tests of Spectral Cloud Classification Using DMSP Fine Mode Satellite Data," AFGL-TR-80-0181, 1980, 42 pp.

R. O. Duda and P. E. Hart, Pattern Classification and Scene Analysis, John Wiley and Sons, Inc., New York, 1973, 482 pp.

R. F. Fournier, "An Initial Study of Power Spectra from Satellite Imagery," AFGL-TR-77-0295, 1977, 28 pp.

Appendix A

RELATIONSHIP BETWEEN THE COMPLEX AND TRIGONOMETRIC FORMS OF THE DISCRETE FOURIER SERIES

Let $\widehat{F}(i,j)$ be a real-valued, discrete function of two variables defined on an $M \times N$ region of $M \cdot N$ equally spaced grid points, where M and N are positive integers. It can be shown that the complex Fourier series of $\widehat{F}$

$$\widehat{F}(i,j) = \sum_{m=\left[-\frac{M}{2}\right]+1}^{\left[\frac{M}{2}\right]} \sum_{n=\left[-\frac{N}{2}\right]+1}^{\left[\frac{N}{2}\right]} \widehat{\psi}_{m,n} e^{2\pi\sqrt{-1}\left(\frac{mi}{M}+\frac{nj}{N}\right)} \;,\; \widehat{\psi}_{m,n} \in \mathbf{C} \;,$$

can be equivalently expressed in trigonometric form

$$\widehat{F}(i,j) = \sum_{m=0}^{\left[\frac{M}{2}\right]} \sum_{n=0}^{\left[\frac{N}{2}\right]} \left\{ \widehat{A}_{m,n} \cos\frac{2\pi mi}{M} \cos\frac{2\pi nj}{N} \right.$$

$$+ \widehat{B}_{m,n} \cos\frac{2\pi mi}{M} \sin\frac{2\pi nj}{N} + \widehat{C}_{m,n} \sin\frac{2\pi mi}{M}$$

$$\left. \cos\frac{2\pi nj}{N} + \widehat{D}_{m,n} \sin\frac{2\pi mi}{M} \sin\frac{2\pi nj}{N} \right\}$$

through use of the following relationships. Let $\widehat{\psi}_{k,1}^{R} = \operatorname{Re}(\widehat{\psi}_{k,1})$, and let $\widehat{\psi}_{k,1}^{I} = \operatorname{Im}(\widehat{\psi}_{k,1})$. Then:

$$\widehat{A}_{0,0} = \widehat{\psi}_{0,0}^{R} \quad (\widehat{\psi}_{0,0}^{I} = 0 \text{ since } \widehat{F} \text{ is real-valued})$$

$$\widehat{B}_{0,0} = \widehat{C}_{0,0} = \widehat{D}_{0,0} = 0$$

$$\widehat{A}_{0,\left[\frac{N}{2}\right]} = \widehat{\psi}_{0,\left[\frac{N}{2}\right]}^{R}$$

$$\widehat{B}_{0,\left[\frac{N}{2}\right]} = \widehat{C}_{0,\left[\frac{N}{2}\right]} = \widehat{D}_{0,\left[\frac{N}{2}\right]} = 0$$

$$\widehat{A}_{\left[\frac{M}{2}\right],0} = \widehat{\psi}_{\left[\frac{M}{2}\right],0}^{R}$$

$$\widehat{B}_{\left[\frac{M}{2}\right],0} = \widehat{C}_{\left[\frac{M}{2}\right],0} = \widehat{D}_{\left[\frac{M}{2}\right],0} = 0$$

$$\widehat{A}_{\left[\frac{M}{2}\right],\left[\frac{N}{2}\right]} = \widehat{\psi}^{R}_{\left[\frac{M}{2}\right],\left[\frac{N}{2}\right]}$$

$$\widehat{B}_{\left[\frac{M}{2}\right],\left[\frac{N}{2}\right]} = \widehat{C}_{\left[\frac{M}{2}\right],\left[\frac{N}{2}\right]} = \widehat{D}_{\left[\frac{M}{2}\right],\left[\frac{N}{2}\right]} = 0$$

$$\widehat{A}_{0,n} = 2\widehat{\psi}^{R}_{0,n}$$

$$\widehat{B}_{0,n} = -2\widehat{\psi}^{I}_{0,n}$$

$$\widehat{C}_{0,n} = \widehat{D}_{0,n} = 0 \qquad\qquad \text{for: } 1 \le n \le \left[\frac{N}{2}\right] - 1$$

$$\widehat{A}_{\left[\frac{M}{2}\right],n} = 2\widehat{\psi}^{R}_{\left[\frac{M}{2}\right],n}$$

$$\widehat{B}_{\left[\frac{M}{2}\right],n} = -2\widehat{\psi}^{I}_{\left[\frac{M}{2}\right],n}$$

$$\widehat{C}_{\left[\frac{M}{2}\right],n} = \widehat{D}_{\left[\frac{M}{2}\right],n} = 0$$

$$\widehat{A}_{m,0} = 2\widehat{\psi}^{R}_{m,0}$$

$$\widehat{C}_{m,0} = -2\widehat{\psi}^{I}_{m,0}$$

$$\widehat{B}_{m,0} = \widehat{D}_{m,0} = 0 \qquad\qquad \text{for: } 1 \le m \le \left[\frac{M}{2}\right] - 1$$

$$\widehat{A}_{m,\left[\frac{N}{2}\right]} = 2\widehat{\psi}^{R}_{m,\left[\frac{N}{2}\right]}$$

$$\widehat{C}_{m,\left[\frac{N}{2}\right]} = -2\widehat{\psi}^{I}_{m,\left[\frac{N}{2}\right]}$$

$$\widehat{B}_{m,\left[\frac{N}{2}\right]} = \widehat{D}_{m,\left[\frac{N}{2}\right]} = 0$$

$$\hat{A}_{m,n} = 2\hat{\psi}^R_{m,n} + 2\hat{\psi}^R_{m,-n}$$

$$\hat{B}_{m,n} = -2\hat{\psi}^I_{m,n} + 2\hat{\psi}^I_{m,-n}$$

$$\hat{C}_{m,n} = -2\hat{\psi}^I_{m,n} - 2\hat{\psi}^I_{m,-n}$$

$$\hat{D}_{m,n} = -2\hat{\psi}^R_{m,n} + 2\hat{\psi}^R_{m,-n}$$

for:

$$1 \leq m \leq \left[\frac{M}{2}\right] - 1, \text{ and}$$

$$1 \leq n \leq \left[\frac{N}{2}\right] - 1$$

Appendix B

PROOF OF THE PERIODICITY OF THE DISCRETE FOURIER TRANSFORM

It is not immediately apparent from equation (6) that
the discrete Fourier transform of the periodic, discrete,
finite function $\hat{F}(i,j)$ (defined in Section 3.1) of two
variables is itself periodic, discrete, and finite. There
are only $M \cdot N$ distinct complex numbers $\hat{\psi}_{m,n}$ which can be com-
puted by (6). Let $m = k$, $n = l$ be any arbitrary integers.
From (6):

$$\hat{\psi}_{k,l} = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \hat{F}(i,j) e^{-2\pi\sqrt{-1}(\frac{ki}{M} + \frac{lj}{N})}.$$

Now, let $m = k+M$, $n = l+N$:

$$\hat{\psi}_{k+M,l+N} = \frac{1}{M \cdot N} \sum_{i=1}^{M} \sum_{j=1}^{N} \hat{F}(i,j) e^{-2\pi\sqrt{-1}\left[\frac{i(k+M)}{M} + \frac{j(l+N)}{N}\right]}$$

$$= \frac{1}{M \cdot N} \sum_{i=1}^{M} \sum_{j=1}^{N} \hat{F}(i,j) e^{-2\pi\sqrt{-1}(\frac{ki}{M} + \frac{lj}{N})}$$

$$e^{-2\pi\sqrt{-1}(\frac{Mi}{M} + \frac{Nj}{N})}$$

$$= \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \hat{F}(i,j) e^{-2\pi\sqrt{-1}(\frac{ki}{M} + \frac{lj}{N})} e^{-2\pi\sqrt{-1}(i + j)}.$$

Using Euler's formula, and since $i + j$ is always an integer,

$$\hat{\psi}_{k+M,l+N} = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \hat{F}(i,j) e^{-2\pi\sqrt{-1}(\frac{ki}{M} + \frac{lj}{N})}.$$

64

$$(\cos 2\pi(i + j) - \sqrt{-1} \sin 2\pi(i + j))$$

$$= \frac{1}{M \cdot N} \sum_{i=1}^{M} \sum_{j=1}^{N} \hat{F}(i,j) e^{-2\pi\sqrt{-1}(\frac{ki}{M} + \frac{lj}{N})}$$

$$= \hat{\psi}_{k,l} \ .$$

Therefore there are only M·N distinct numbers for which (6) can be evaluated. Thus, if the function $\hat{F}(i,j)$ defined for an M x N grid is one complete cycle of a discrete, periodic function, then its Fourier transform is also one complete cycle of a discrete, periodic function with M·N unique coefficients in each period.

Appendix C

THE DFT OF A REAL-VALUED FUNCTION

In order for the complex Fourier series expansion of a real-valued, discrete, finite function $\hat{F}(i,j)$ to be real-valued, the following relationships must hold:

$$\hat{\psi}_{m,n} = \hat{\psi}^{*}_{-m,-n}$$

$$\hat{\psi}_{m,-n} = \hat{\psi}^{*}_{-m,n}$$

where $\hat{\psi}^{*}_{m,n}$ denotes the complex conjugate of $\hat{\psi}_{m,n}$. These relationships can be derived by expanding (5) and setting the coefficients of all the resultant imaginary terms equal to zero. If $\hat{\psi}_{m,n} = \hat{\psi}^{*}_{-m,-n}$, then $P_{m,n} = P_{-m,-n}$ since (from (9))

$$P_{-m,-n} = \sqrt{(\hat{\psi}^{R}_{-m,-n})^2 + (\hat{\psi}^{Im}_{-m,-n})^2}$$

$$= \sqrt{(\hat{\psi}^{R}_{m,n})^2 + (-\hat{\psi}^{Im}_{m,n})^2}$$

$$= \sqrt{(\hat{\psi}^{R}_{m,n})^2 + (\hat{\psi}^{Im}_{m,n})^2}$$

$$= P_{m,n} \ .$$

Thus by using only the powers of those coefficients which lie in the first quadrant of the m-n plane (as does the "first quadrant" method described in the Results section), any information contained in the powers of those coefficients which lie in the third quadrant is used as well. Essentially one-half the information contained in the transform is utilized when one-fourth the available coefficients of that transform are processed.

Appendix D

CLASSIFIER ACCURACY MATRICES

The following Figures D-1 through D-5 contain matrices which present the results of the automated classifiers for the set of 143 cloud imagery samples whose composition is shown in Figure 7. Each figure is labeled in order to indicate which classifier variation was utilized to obtain its shown results. The cloud sample classifications made by skilled analysts are represented by the rows of each matrix (TRUTH), and the decisions made by the automated classifier are represented by the columns of each matrix (CLASSIFIER). Agreement between the classifier and truth is shown along the main diagonal (upper left to lower right) of each matrix. Each pair of matrices contains a frequency distribution on the left, and its corresponding relative frequency distribution on the right. The figures show in detail exactly how the automated classifiers performed.

Overall relative frequency for all or any group of clouds is computed by summing the number of correct hits within the cloud type or cloud group of interest, and then dividing that sum by the total number of samples within that cloud type or group.

CLASSIFIER — VISIBLE

VISIBLE TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	2		1			2		1	1
St	1	5		2	1				1
Cu	3		10	9	2				
CLEAR	1		6	8	2		2		
Ac	1	3		1	1		2	1	
As/Ns					1	3			1
Ci	8	1	6	4	3	3	9	2	
Cs	1	1	2			1	4	9	1
Cb					2	3	2	3	4

VISIBLE TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	.29		.14			29		.14	.14
St	.10	.50		.20	.10				.10
Cu	.13		.42	.38	.07				
CLEAR	.05		.32	.42	.11		.11		
Ac	.11	.33		.11	.11		.22	.11	
As/Ns					.20	.60			.20
Ci	.22	.03	.17	.11	.08	.08	.25	.06	
Cs	.05	.05	.11			.05	.21	.47	.05
Cb					.14	.21	.14	.21	.29

CLASSIFIER — INFRARED

INFRARED TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	2	1	1			2	1		
St	2	2		1			1	4	
Cu	2	3	14	5					
CLEAR			1	4	14				
Ac	1	2				2	2	1	1
As/Ns		1			1	2	1		
Ci	5	5	5	4		6	6		5
Cs			1	1				8	9
Cb	1					1	1	2	9

INFRARED TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	.29	.14	.14			.29	.14		
St	.20	.20		.10			.10	.40	
Cu	.08	.13	.58	.21					
CLEAR			.05	.21	.74				
Ac	.11	.22				.22	.22	.11	.11
As/Ns		.20			.20	.40	.20		
Ci	.14	.14	.14	.11		.17	.17		.14
Cs			.05	.05				.42	.47
Cb	.07					.07	.07	.14	.64

CLASSIFIER — COMBINED VIS & IR

COMBINED VIS & IR TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	2	1	1			2	1		
St	1	6		1	1	1			
Cu	2	2	14	5	1				
CLEAR			7	11			1		
Ac	1	2			4		1	1	
As/Ns		1			1	3			
Ci	5	3	4		3	3	15	1	2
Cs		2				1	1	12	3
Cb		1				2	2	2	7

COMBINED VIS & IR TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	.29	.14	.14			.29	.14		
St	.10	.60		.10	.10	.10			
Cu	.08	.08	.58	.21	.04				
CLEAR			37	.58			.05		
Ac	.11	.22			.44		.11	.11	
As/Ns		.20			.20	.60			
Ci	.14	.08	.11		.08	.08	.42	.03	.06
Cs		.11				.05	.05	.63	.16
Cb		.07				.14	.14	.14	.50

Figure D-1. Mean Classifier.

67

CLASSIFIER — VISIBLE TRUTH (counts)

TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	4		1				2		
St	1	4					3		2
Cu	3		16				3		2
CLEAR				14			3	2	
Ac	1		2		2		3		1
As/Ns						1	3		1
Ci	4	3		4	1		15	5	4
Cs				1			4	14	
Cb			1				1	1	11

CLASSIFIER — VISIBLE TRUTH (fractions)

TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	.57		.14				.29		
St	.10	.40					.30		.20
Cu	.13		.67				.13		.08
CLEAR				.74			.16	.11	
Ac	.11		.22		.22		.33		.11
As/Ns						.20	.60		.20
Ci	.11	.08		.11	.03		.42	.14	.11
Cs			.05				.21	.74	
Cb			.07				.07	.07	.79

CLASSIFIER — INFRARED TRUTH (counts)

TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	2		2	1			2		
St	1	2		3			2	2	
Cu	2		13		1		4	4	
CLEAR	1		2	14			2		
Ac	1			1	3		2	2	
As/Ns				3			1	1	
Ci	3		3				26	2	2
Cs	1						2	16	
Cb							2	1	11

CLASSIFIER — INFRARED TRUTH (fractions)

TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	.29		.29	.14			.29		
St	.10	.20		.30			.20	.20	
Cu	.08		.54		.04		.17	.17	
CLEAR	.05		.11	.74			.11		
Ac	.11			.11	.33		.22	.22	
As/Ns				.60			.20	.20	
Ci	.08		.08				.72	.06	.06
Cs	.05						.11	.84	
Cb							.14	.07	.79

CLASSIFIER — COMBINED VIS & IR TRUTH (counts)

TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	6						1		
St	1	7					2		
Cu		1	19		1		1	1	1
CLEAR	2			14			1	2	
Ac	1		1		6		1		
As/Ns						5			
Ci	2	1	1				30		2
Cs							2	17	
Cb							1	1	12

CLASSIFIER — COMBINED VIS & IR TRUTH (fractions)

TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	.86						.14		
St	.10	.70					.20		
Cu		.04	.79		.04		.04	.04	.04
CLEAR	.11			.74			.05	.11	
Ac	.11		.11		.67		.11		
As/Ns						1.0			
Ci	.06	.03	.03				.83		.06
Cs							.11	.89	
Cb							.07	.07	.86

Figure D-2. Spectral classifier without a priori and co-variance terms, "all-coefficient" method.

VISIBLE — CLASSIFIER (counts)

TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	3		1	1			1		1
St		7					1		2
Cu	2		14		3		2	1	2
CLEAR				18			1		
Ac			1		7		1		
As/Ns						5			
Ci		3		5	1		16	8	3
Cs				1	1		1	16	
Cb					1		1	1	11

VISIBLE — CLASSIFIER (coefficients)

TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	.43		.14	.14			.14		.14
St		.70					.10		.20
Cu	.08		.58		.13		.08	.04	.08
CLEAR				.95				.05	
Ac			.11		.78		.11		
As/Ns						1.0			
Ci		.08		.14	.03		.44	.22	.08
Cs			.05	.05			.05	.84	
Cb					.07		.07	.07	.79

INFRARED — CLASSIFIER (counts)

TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	3	1			1			2	
St		8	1		1				
Cu	7	1	12				3	1	
CLEAR	1			17				1	
Ac	1			1	4			2	1
As/Ns						5			
Ci	4	2	1				26	1	2
Cs	1		2		1			14	1
Cb					2		1	1	10

INFRARED — CLASSIFIER (coefficients)

TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	.43	.14			.14			.29	
St		.80	.10		.10				
Cu	.29	.04	.50				.13	.04	
CLEAR	.05			.89				.05	
Ac	.11			.11	.44			.22	.11
As/Ns						1.0			
Ci	.11	.06	.03				.72	.03	.06
Cs	.05		.11		.05			.74	.05
Cb					.14		.07	.07	.71

COMBINED VIS & IR — CLASSIFIER (counts)

TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	5						1	1	
St		10							
Cu	2	1	17		1		1	1	1
CLEAR	1			17				1	
Ac	2				7				
As/Ns						5			
Ci	2	2	1				23	6	2
Cs					1		1	17	
Cb							1	1	12

COMBINED VIS & IR — CLASSIFIER (coefficients)

TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	.71						.14	.14	
St		1.0							
Cu	.08	.04	.71		.04		.04	.04	.04
CLEAR	.05			.89				.05	
Ac	.22				.78				
As/Ns						1.0			
Ci	.06	.06	.03				.64	.17	.06
Cs					.05		.05	.89	
Cb							.07	.07	.86

Figure D-3. Spectral classifier with a priori and covariance terms, "all-coefficient" method.

69

VISIBLE

CLASSIFIER

TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	4		1				2		
St	1	4					3	2	
Cu	2		15			.1	3		3
CLEAR				14			4	1	
Ac			1		3		4		1
As/Ns	1						2		2
Ci	6	1		5			20	2	2
Cs				1			4	14	
Cb							1	1	12

CLASSIFIER

TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	.57		.14				.29		
St	.10	.40					.30		.20
Cu	.08		.63			.04	.13		.13
CLEAR				.74			.21	.05	
Ac			.11		.33		.44		.11
As/Ns	.20						.40		.40
Ci	.17	.03		.14			.56	.06	.06
Cs			.05				.21	.74	
Cb							.07	.07	.86

INFRARED

CLASSIFIER

TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	2		1	1			1	2	
St		2		3			2	3	
Cu	2		17		1		3	1	
CLEAR			3	15	1				
Ac			1		5			1	2
As/Ns				3			1	1	
Ci	2		4	·			26	1	3
Cs	1		1				2	15	
Cb					2		1	1	10

CLASSIFIER

TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	.29		.14	.14			.14	.29	
St		.20		.30			.20	.30	
Cu	.08		.71		.04		.13	.04	
CLEAR			.16	.79	.04				
Ac			.11		.56			.11	.22
As/Ns				.60			.20	.20	
Ci	.06		.11				.72	.03	.06
Cs	.04		.04				.11	.79	
Cb					.14		.07	.07	.71

COMBINED VIS & IR

CLASSIFIER

TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	5		1				1		
St		8					2		
Cu	1		19		1			1	2
CLEAR				15			3	1	
Ac	1		1		5		2		
As/Ns					1	4			
Ci	3		1				30	2	
Cs							2	17	
Cb							1	1	12

CLASSIFIER

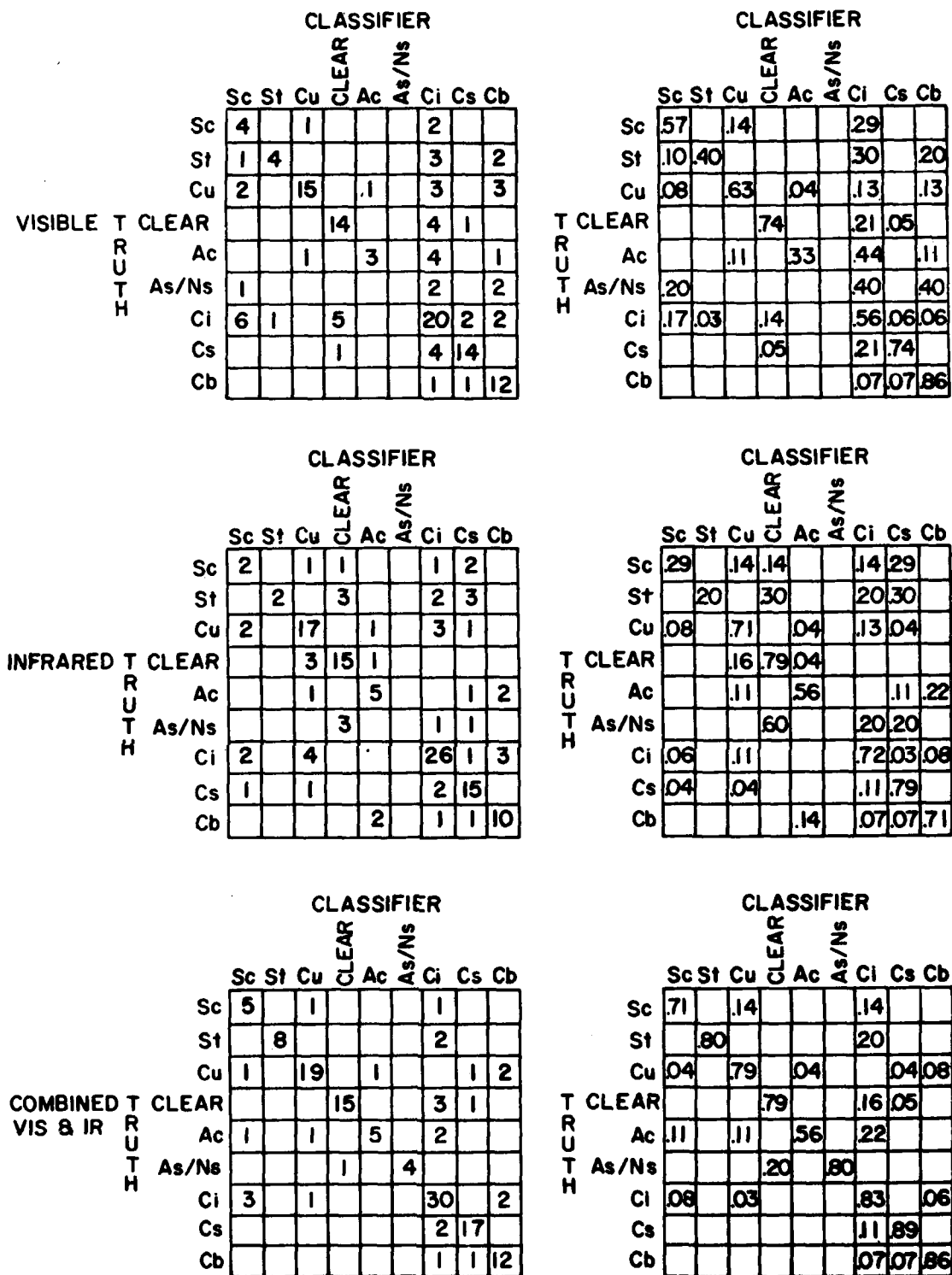| TRUTH	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	.71		.14				.14		
St		.80					.20		
Cu	.04		.79		.04			.04	.08
CLEAR				.79			.16	.05	
Ac	.11		.11		.56		.22		
As/Ns					.20	.80			
Ci	.08		.03				.83	.06	
Cs							.11	.89	
Cb							.07	.07	.86

Figure D-4. Spectral classifier without a priori and covariance terms, "quarter-coefficient" method.

CLASSIFIER (counts) — VISIBLE TRUTH

TRUTH \ CLASSIFIER	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	3		1	1			1		1
St		7						1	2
Cu			18		1		2	1	2
CLEAR				18				1	
Ac			1		7				1
As/Ns						5			
Ci	1	1		5			21	6	2
Cs				1			1	17	
Cb							1	1	12

CLASSIFIER (probabilities) — VISIBLE TRUTH

TRUTH \ CLASSIFIER	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	.43		.14	.14			.14		.14
St		.70						.10	.20
Cu			.75		.04		.08	.04	.08
CLEAR				.95				.05	
Ac			.11		.78				.11
As/Ns						1.0			
Ci	.03	.03		.14			.58	.17	.06
Cs				.05			.05	.89	
Cb							.07	.07	.86

CLASSIFIER (counts) — INFRARED TRUTH

TRUTH \ CLASSIFIER	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	3			1				3	
St		7	1					1	1
Cu	7		11	1			4	1	
CLEAR	1			18					
Ac	1			1	2			3	2
As/Ns						5			
Ci	4	3					25	1	3
Cs			2				2	15	
Cb					2		1	1	10

CLASSIFIER (probabilities) — INFRARED TRUTH

TRUTH \ CLASSIFIER	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	.43			.14				.43	
St		.70	.10					.10	.10
Cu	.29		.46	.04			.17	.04	
CLEAR	.05			.95					
Ac	.11			.11	.22			.33	.22
As/Ns						1.0			
Ci	.11	.08					.69	.03	.08
Cs			.11				.11	.79	
Cb					.14		.07	.07	.71

CLASSIFIER (counts) — COMBINED VIS & IR TRUTH

TRUTH \ CLASSIFIER	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	5		1				1		
St		10							
Cu	2		19					1	2
CLEAR				17			1	1	
Ac	1		2		6				
As/Ns						5			
Ci	2		1				26	5	2
Cs							2	17	
Cb							1	1	12

CLASSIFIER (probabilities) — COMBINED VIS & IR TRUTH

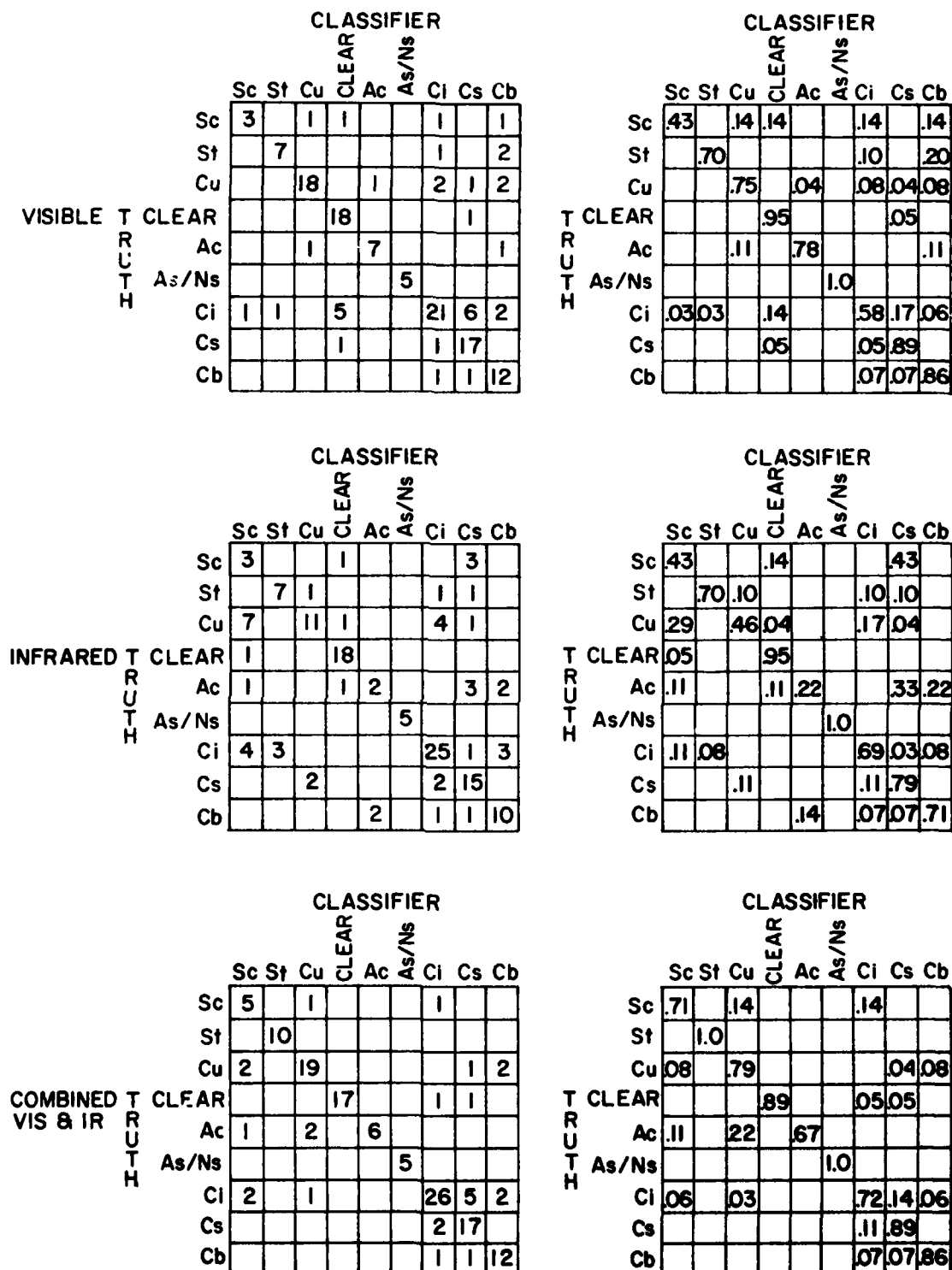| TRUTH \ CLASSIFIER	Sc	St	Cu	CLEAR	Ac	As/Ns	Ci	Cs	Cb
Sc	.71		.14				.14		
St		1.0							
Cu	.08		.79					.04	.08
CLEAR				.89			.05	.05	
Ac	.11		.22		.67				
As/Ns						1.0			
Ci	.06		.03				.72	.14	.06
Cs							.11	.89	
Cb							.07	.07	.86

Figure D-5. Spectral classifier with a priori and covariance terms, "quarter-coefficient" method.

71

Appendix E

CLASSIFICATION USING A TRI-DIAGONAL COVARIANCE MATRIX

The combined visible/infrared classifier used in this study computed the 26-component power spectrum for the visible and the IR of each 37 x 37-pixel imagery sample it classified. Pickett and Blackman (1979) suggested a different method of combining the two channels.

Instead of a 26-component visible (**x**) and infrared power spectrum (**y**), this method requires that a 2·26 = 52-component visible/infrared spectrum **z** be used as input to the discriminant function (16). The first component of this spectrum is the mean visible value NAA_0^{vis} of an imagery sample, followed by that 37 x 37 sample's 25 visible spectral components $NAA_n^{vis} \equiv x_n$, n = 1,2,3,...,25. The 27th component of **z** is

the mean IR value NAA_0^{IR} of the sample, followed by 25 infra-

red spectral components $NAA_m^{IR} \equiv y_m$, m = 1,2,3,...,25.

The original 26 x 26 diagonal covariance matrix $\sum_i$ would increase in size to a 52 x 52 tri-diagonal covariance matrix $\sum_i'$, composed of four smaller 26 x 26 submatrices, all of which are diagonal. The main diagonals of the upper left and lower right submatrices consist of the variances of the visible and the variances of the IR spectral components whose annular band numbers are the same. The main diagonals of the lower left and the upper right submatrices consist of the covariances of the visible and infrared spectral components whose annular band numbers are the same.

The 52-component column vector **z** and the larger tri-diagonal covariance matrix $\sum_i'$ are then plugged into equation (16), with r = 52. Thus for a 37 x 37 cloud imagery sample input array,

$$\mathbf{z} = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_{25} \\ y_0 \\ y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_{25} \end{bmatrix} = \begin{bmatrix} NAA_0^{vis} \\ NAA_1^{vis} \\ NAA_2^{vis} \\ \cdot \\ \cdot \\ \cdot \\ NAA_{25}^{vis} \\ NAA_0^{IR} \\ NAA_1^{IR} \\ NAA_2^{IR} \\ \cdot \\ \cdot \\ \cdot \\ NAA_{25}^{IR} \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ \cdot \\ \cdot \\ \cdot \\ z_{26} \\ z_{27} \\ z_{28} \\ z_{29} \\ \cdot \\ \cdot \\ \cdot \\ z_{52} \end{bmatrix} ,$$

and

$$\sum_i' = \begin{bmatrix} \sum_i^{vis} & \vdots & \sum_i^{vis,IR} \\ \cdots\cdots\cdots & \vdots & \cdots\cdots\cdots \\ \sum_i^{vis,IR} & \vdots & \sum_i^{IR} \end{bmatrix} ,$$

so that (from (16))

$$d_i(\mathbf{z}) = -\frac{1}{2}(\mathbf{z} - \boldsymbol{\mu}_i')^t (\sum_i')^{-1}(\mathbf{z} - \boldsymbol{\mu}_i') - \frac{52}{2}\ln 2\pi - \frac{1}{2}\ln \left| \sum_i' \right|$$

$$+ \ln P(\omega_i), \quad i = 1,2,3,\ldots,9.$$

The column vector $\boldsymbol{\mu}_i'$ is the mean vector of all vectors $\mathbf{z}$ generated from samples containing cloud type i. The decision rule would choose class k if $d_k > d_i$, for all $i \neq k$.

GLOSSARY OF TERMS

DMSP	Defense Meteorological Satellite Program
DFT	discrete Fourier transform
$\mathbf{R}$	the set of all real numbers
$\mathbf{R}^2$	the set of all ordered pairs (x,y) such that x and y are real numbers
$\mathbf{C}$	the set of all complex numbers
$\mathbf{Z}$	the set of all integers
$\mathbf{Z} \times \mathbf{Z}$	the set of all points (i,j) such that i and j are integers
$\forall$...	for every...
...$\in$...	...is an element of...
$[x]$	the greatest integer $\leq x$, x positive; $[-x] = -[x]$
$\hat{\mathbf{F}}(i,j)$	a discrete, finite approximation to the continuous function $\mathbf{F}(x,y)$
FFT	fast Fourier transform
$Re(z), Im(z), z^*$	Let $z = a + bi$ be a complex number, $i = \sqrt{-1}$. Then: $Re(z) = a$ $Im(z) = b$ $z^* = a - bi$ (complex conjugate of z)
$\mathbf{x}$	power spectrum of a visible imagery sample
$\mathbf{y}$	power spectrum of an infrared imagery sample
Sc	stratocumulus
St	stratus
Cu	cumulus
Ac	altocumulus
As/Ns	altostratus/nimbostratus
Ci	cirrus
Cs	cirrostratus
Cb	cumulonimbus
Euler's formula	$e^{i\theta} = \cos\theta + i\sin\theta$, $i = \sqrt{-1}$
$|\Sigma|$	determinant of the square matrix Σ