

N O T I C E

THIS DOCUMENT HAS BEEN REPRODUCED FROM
MICROFICHE. ALTHOUGH IT IS RECOGNIZED THAT
CERTAIN PORTIONS ARE ILLEGIBLE, IT IS BEING RELEASED
IN THE INTEREST OF MAKING AVAILABLE AS MUCH
INFORMATION AS POSSIBLE

AgRISTARS

SR-LO-00418 80-10119
JSC-13776 NASA CR

160541

Available under NASA sponsorship
in the interest of early and wide dis-
tribution of Earth Resources Survey
information and without liability
for use made thereof.

A Joint Program for
Agriculture and
Resources Inventory
Surveys Through
Aerospace
Remote Sensing

Supporting Research

February 1980

TECHNICAL REPORT

INTERPRETATION OF LANDSAT DIGITAL DATA USING
A CUBIC COLOR MODEL BASED ON
RELATIVE ENERGIES

R. B. Cate, D. E. Phinney, M. C. Kinsler,
M. L. Sestak, T. Hodges, and J. J. Dishler

(E80-10119) INTERPRETATION OF LANDSAT
DIGITAL DATA USING A CUBIC COLOR MODEL BASED
ON RELATIVE ENERGIES (Lockheed Engineering
and Management) 31 p HC A03/MF A01 CSCL 05B

N80-23745

Unclas
G3/43 00119



NASA



LOCKHEED ENGINEERING AND MANAGEMENT SERVICES COMPANY, INC.
1830 NASA Road 1, Houston, Texas 77058

SR-LO-00418
JSC-13776

TECHNICAL REPORT
INTERPRETATION OF LANDSAT DIGITAL DATA USING
A CUBIC COLOR MODEL BASED ON
RELATIVE ENERGIES


Job Order 73-312

This report describes Vegetation/Soils/Field Research activities
of the Supporting Research Project of the AgRISTARS program.


Prepared By

R. B. Cate, D. E. Phinney, M. C. Kinsler,
M. L. Sestak, T. Hodges, and J. J. Dishler

APPROVED BY



D. E. Phinney, Supervisor
Agriculture Technology Section



J. E. Wainwright, Manager
Development and Evaluation Department

LOCKHEED ENGINEERING AND MANAGEMENT SERVICES COMPANY, INC.
Under Contract NAS 9-15800

For

Earth Observations Division
Space and Life Sciences Directorate

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

February 1980

LEMSCO-13499

PRECEDING PAGE BLANK NOT FILMED

CONTENTS

Section	Page
1. INTRODUCTION	1-1
2. SCENE STANDARDIZATION.	2-1
3. COLOR TRANSFORMATION	3-1
4. SUMMARY AND CONCLUSION	4-1
5. BIBLIOGRAPHY	5-1
Appendix	
A. CORRELATION STRUCTURE IN THE NORMALIZED DATA	A-1

PRECEDING PAGE BLANK NOT FILMED

TABLES

Table		Page
1	SOURCES OF VARIATION IN MEANS OF LANDSAT DATA ACQUISITIONS	2-2
2	CORRELATION STRUCTURE OF THE LANDSAT DIGITAL DATA FOR 18 550 RANDOM SAMPLES OF SPRING WHEAT SCENES	2-8
3	GENERALIZED COLOR SEQUENCE FOR AGRICULTURAL CROPS.	3-9

FIGURES

Figure		Page
1	The relationship between scene means and dots selected randomly from spring wheat scenes, Landsat channel 1	2-3
2	The relationship between scene means and dots selected randomly from spring wheat scenes, Landsat channel 2	2-4
3	The relationship between scene means and dots selected randomly from spring wheat scenes, Landsat channel 3	2-5
4	The relationship between scene means and dots selected randomly from spring wheat scenes, Landsat channel 4	2-6
5	Color normalization and reproduction using relative energy (RE) concept.	3-3
6	Hue distribution in the color cube	3-4
7	Relationship of hue, value, and chroma notation to the color cube	3-5
8	Diagram of color cube hue notation	3-6
9	Channel ranking hue domains of the color cube.	3-8

PRECEDING PAGE BLANK NOT FILMED

1. INTRODUCTION

The availability of large quantities of digital (numerical) multispectral scanner data (MSS) from the Land Satellite (Landsat) system has contributed greatly to the rapid development of remote sensing technology. Parallel evolution of machine processing techniques for these data has resulted from desires for greater accuracy and for the ability to monitor larger and larger areas. One of the major beneficiaries of these developments has been agriculture. Remote sensing provides a unique data source for monitoring food and fiber production on a worldwide scale.

At the heart of such applications systems lies the processing and presentation of the basic digital data. Effective, yet simple, isolation of the informational content of such data is required. The digital data are frequently transformed to enhance features of particular interest and used either directly in numerical algorithms or through image interpretation of film products. Quantitative estimates of plant characteristics such as yield, quality, and growth stage have frequently relied on direct use of digital data through statistically derived predictive equations. Crop identification and acreage estimation generally have involved the use of numerical pattern recognition in conjunction with visual interpretation of film products. Typically, training data for supervised classification algorithms have been provided through the identification and labeling of homogeneous areas in photographic products. Generally, the multitemporal progression of colors of a given field or fields is used for labeling purposes. The colors are determined by an analyst from a false color-infrared (IR) film product using Landsat bands 4, 5, and 7.¹

The purpose of this paper is to present a proposed transformation between digital data and color space. The underlying assumption is that transforming multispectral data into color notation provides a powerful analysis

¹Henceforth, in this paper we will refer to Landsat MSS bands 4 through 7 as channels 1 through 4.

tool for development of objective techniques useful in both image interpretation and numerical modeling. In other words, color can serve as the primary mechanism for extracting information from Landsat data. The previously mentioned use of multitemporal color sequences in labeling is an obvious example. A stable transform between digital data and film products would allow for identification, verification, and utilization of subtle variations in color. These are normally difficult to quantify, yet they supply essential information in the decision logic used by experienced image interpreters. This transform would reduce the subjective nature of the analyst decisions that makes results difficult to replicate. Ultimately, the movement from research and development to general usage would become a transfer of technology rather than of experience and art. In addition, a substantial body of agricultural literature exists which relates plant characteristics such as stress (moisture and nutrient), quality and key morphological changes to color.

Significant variations, both within scenes over time and between scenes, produced by atmospheric radiative processes have hindered the quantitative interpretation of Landsat digital data. Radiative transfer models have been used to identify and evaluate the magnitude of the sources of variation (Lambeck and Potter, 1979). However, application of these models for purposes of correcting the Landsat signal has been impractical due to the operational shortage of data on key parameters such as total optical depth. To date, attempts to solve this problem in application have met with only limited success. Sufficient variability has remained such that it has only been feasible to classify Landsat measurements within individual scenes. For similar reasons, algorithms to estimate plant characteristics have shown a tendency to be dominated by gross location-to-location variations and be relatively insensitive to within-scene variations.

The transform discussed in the following sections is intended to provide a stable transformation between digital data and color space which maximizes the spatial and temporal comparability of Landsat data. The color space is both logical in terms of color theory and compatible with the hardware systems used to produce color film products.

2. SCENE STANDARDIZATION

The spatial and temporal variability of agricultural scenes is illustrated in figures 1 through 4, which are plots of 18 550 random picture elements (pixels) sampled from 371 Landsat acquisitions from 35 LACIE spring wheat ground-truth sites during the spring, summer, and fall of 1978. The pixel values are plotted against the means of the acquisitions from which they were sampled. These figures show that about half of the variability among pixel channel measurements can be accounted for by variability in the means of the acquisitions in which they occur. This is a one to one relationship, with a slope of one and an intercept of zero. The variability in the means of the acquisitions can be caused by both scene-wide atmospheric effects (illumination, transmission, scattering, calibration, noise) and differences in scene composition (proportions of plants and soils of varying appearance). Such sources of variation are shown in table 1. These are difficult to estimate separately, but the combined multiplicative effect can be neutralized approximately if all pixel channel values are divided by the channel mean for the acquisition. However, some of the sources of variation (especially scattering and noise) have an additive effect and tend to mask the signal. Hence, the utility of the channel mean preprocessing may be reduced as atmospheric haze increases. A subsequent paper will discuss the results of simulation and statistical studies designed to elucidate further the extent to which the efficiency of the mean standardization technique is affected by the various components of acquisition variability. Details of the proposed preprocessing technique are presented in the next paragraph.

With the assumption that the additive and multiplicative terms are the same for both the scene and a element within the scene, the standardization may be written as follows:

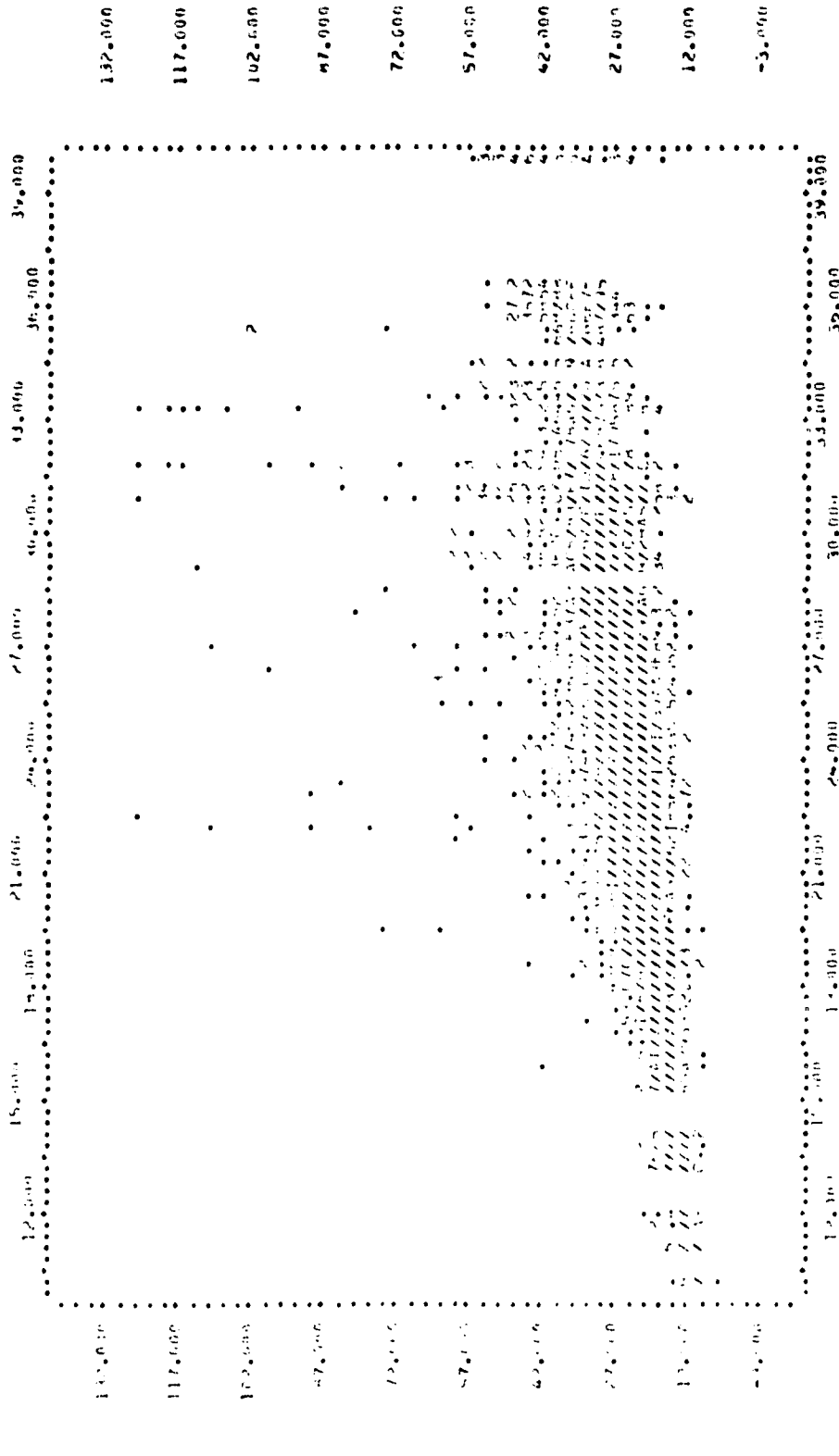
$$RE_i = \frac{m_i S_i (\text{element}) + a_i}{m_i S_i (\text{scene}) + a_i}$$

where the standardized data are expressed as relative energy (RE), S_i represents the measured spectral signal for the i^{th} Landsat band and m_i and a_i are multiplicative and additive terms, respectively.

TABLE 1.- SOURCES OF VARIATION IN MEANS OF LANDSAT DATA ACQUISITIONS

Variable	Source	Effect type	
		Multiplicative	Additive
Illumination	Sun angle, atmospheric scattering, path transmission	X	
Transmission	Zenith path transmission, atmospheric scattering	X	
Path radiance	Atmospheric backscatter		X
Calibration	Electronics degradation	X	
Noise	System components		X
Scene composition	Differences in amounts of soils and plants of varying appearance	X	

TSS (OT Data)
 PLOT OF SEG. CH1 (LANDSAT) VS. SEG. CH1 (TSS-ARIS).
 10000 points and 100000



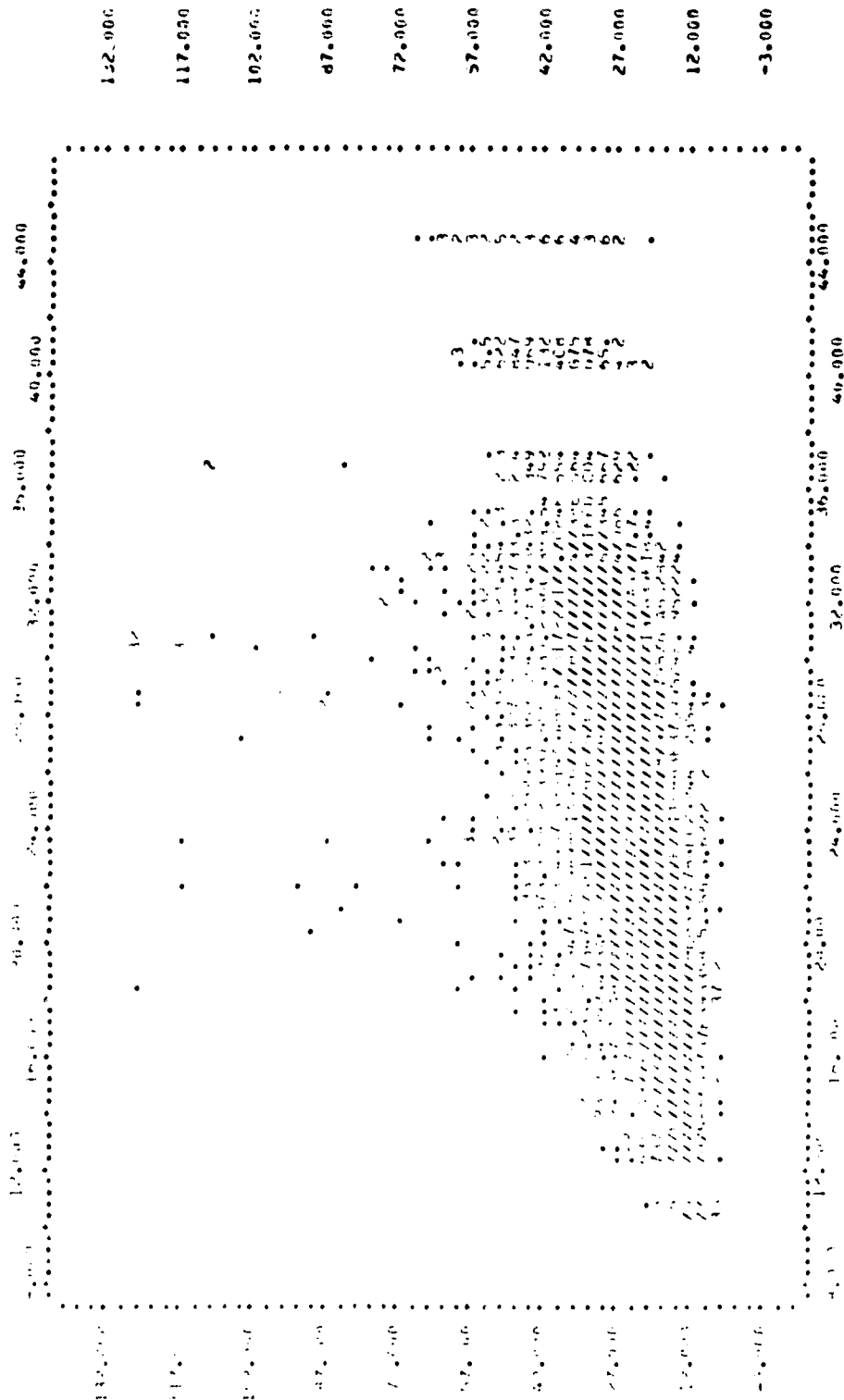
SIMPLE STATISTICS:

VARIABLE	MEAN	STD	RSQ	INTERCEPT	SLOPE
SEG CH1	23.8074	4.9762	0.4610	0.7008	0.9727
RAW CH1	23.8574	7.1286			

Figure 1.- The relationship between scene means and dots selected randomly from spring wheat scenes, Landsat channel 1.

2-3
5

TVS JOB CMT
 PLOT OF Y VS X (X=RAW CHZ (T-AXIS),
 Y=SEG CHZ (R-AXIS))



SIMPLE STATISTICS:

VARIABLE	MEAN	STD	RSQ	INTERCEPT	SLOPE
SEG CHZ	23.4847	6.0999	0.4146	0.6278	0.9737
RAW CHZ	23.4960	9.2251			

Figure 2.- The relationship between scene means and dots selected randomly from spring wheat scenes, Landsat channel 2.

2-4

TYS: 001 071A
 PAGE OF 500 (44-151) 000, 000, 000 (Y-0000)
 1000 0000 0000 0000 0000



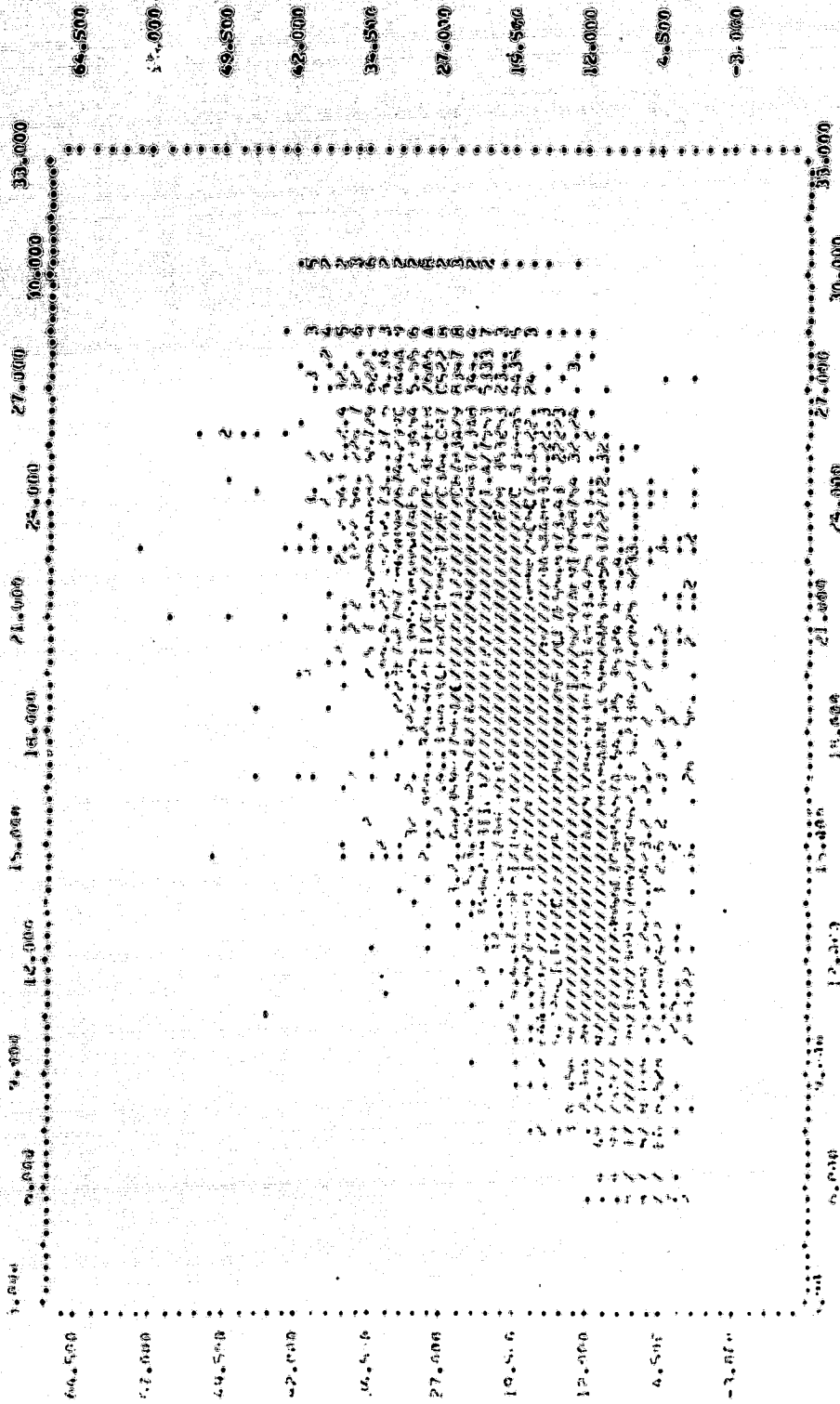
SIMPLE STATISTICS:
 VARIABLE MEAN STD RSQ INTERCEPT SLOPE
 SEG CH3 38.1168 9.5092 0.5286 0.0756 1.0031
 RAW CH3 38.3102 13.1194

Figure 3.- The relationship between scene means and dots selected randomly from spring wheat scenes, Landsat channel 3.

ORIGINAL PAGE IS
 OF POOR QUALITY

2-5
 7

TYPE POINT DATA
 POINT OF SEG CH4 (4-11-15) VS. CH4 (4-11-15)
 10000 POINTS ARE PLOTTED.



SIMPLE STATISTICS:

VARIABLE	MEAN	STD	RSQ	INTERCEPT	SLOPE
SEG CH4	17.7956	4.6815	0.5158	0.2297	0.9935
RAW CH4	17.9100	6.4761			

Figure 4. - The relationship between scene means and dots selected randomly from spring wheat scenes, Landsat channel 4.

26
8

Alternatively, the standardization can be expressed as follows:

1. Scene channel means (CH_i) are individually standardized to a constant, K, so that $K/CH_i =$ channel correction factor (CCF_i)
2. Each digital value, by channel, is then multiplied by the appropriate correction factor to obtain a relative energy (RE) value

$$RE_i = CCF_i * \text{raw value}$$

For use in numerical algorithms, we have chosen to assign K the arbitrary value of 5, corresponding to the midpoint of a 0 to 10 relative energy scale. This practice provides continuity with other colorimetric systems, such as the Munsell (1963), which place brightness on a 0 to 10 scale. A value of $K = 128$ is used in the creation of false color imagery as discussed in the next section.

It can be shown mathematically that the data structure (see the derivation in the appendix²) is unchanged by this transformation. Table 2 gives empirical verification of this assertion, using the 18 550 observations plotted in figures 1 through 4. The table shows the between-channel correlation matrix for both the raw digital data and the transformed values. Clearly, the interchannel correlation is not materially affected by a transformation of this type.

²The authors would like to acknowledge the aid of Dr. R. S. Chhikara in preparing the statistical analyses of the correlation properties of the scene normalization procedure, as presented in the appendix of this document.

TABLE 2.- CORRELATION STRUCTURE OF THE LANDSAT DIGITAL DATA
FOR 18 550-RANDOM SAMPLES OF SPRING WHEAT SCENES

(a) Raw data

	<u>Channel 1</u>	<u>Channel 2</u>	<u>Channel 3</u>	<u>Channel 4</u>
Channel 1	1.000			
Channel 2	.920	1.000		
Channel 3	.520	.352	1.000	
Channel 4	.301	.113	.940	1.000

(b) Relative energy values

	<u>Channel 1</u>	<u>Channel 2</u>	<u>Channel 3</u>	<u>Channel 4</u>
Channel 1	1.000			
Channel 2	.906	1.000		
Channel 3	.473	.351	1.000	
Channel 4	.286	.134	.937	1.000

~~28~~
10

3. COLOR TRANSFORMATION

In most classical theory, color is held to be determined by the absolute amounts of radiation in the three primary bands or channels of green, blue, and red. However, E. H. Land and his colleagues at Polaroid have amassed considerable evidence that human perception of repeatable color is better correlated with relative amounts of radiation rather than with absolute radiation (Land, 1974, 1977). Land's definition of relative varies but is approximately equivalent to the percentage of incident radiation. However, the behavior of the Landsat data illustrated in figures 1 through 4 suggests that a better scientific color reference point is the normalized scene mean.

The proposed standardization procedure provides numbers that are relative to a neutral mean in which all channel means are equal to five. Our hypothesis is that all colors can be defined or classified by their positions within a cube whose three primary color axes (blue, green and red) have been scaled from zero to ten in these standardized numbers, called relative energies. A logical corollary is that human perception of color is analogous to this relative energy standardization and classification.

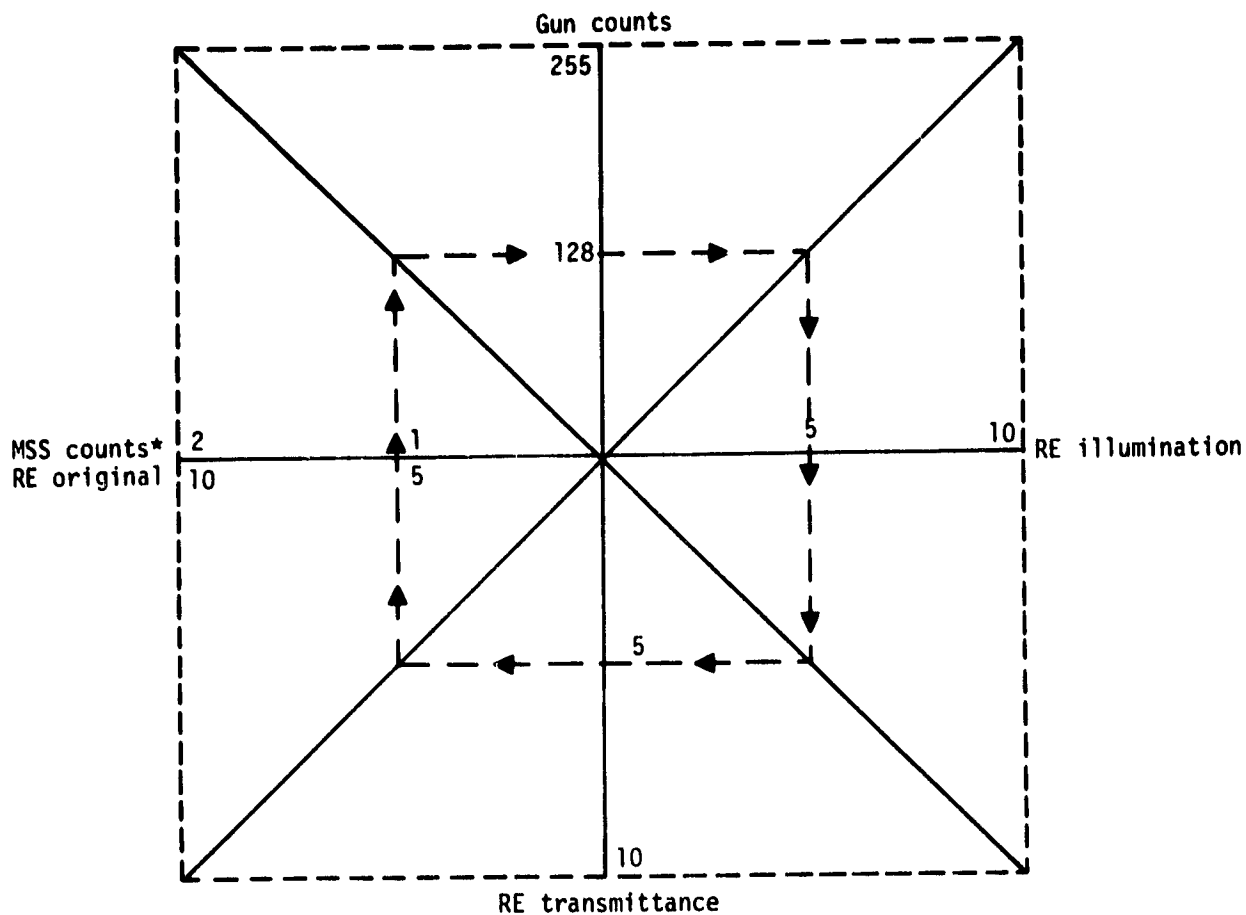
There is considerable empirical evidence for the validity of the above hypothesis. Indeed, it seems to explain the effectiveness of a whole class of devices used to convert digital data to color representations on either film or cathode ray tube (CRT) screens. One such device is the Production Film Converter (PFC). It consists of a CRT light source with three filters, blue, green, and red. The amount of illumination through each filter for each pixel is controlled by Landsat digital tapes. Illumination is expressed in gun-counts and these range from 0 to 255. The machine is calibrated so that an illumination of 255 gun-counts in each channel will just produce colorless transparency (white) when the film product is viewed on a standard light table. The Landsat digital data are converted to gun-counts in a manner which results in the scene mean of each channel being equal to 128 gun-counts. This is a balancing technique to ensure that the scene mean is halfway between maximum opacity (black, created by zero gun-counts in each channel) and maximum transparency (white, created by 255 gun-counts in each channel).

Thus, it is a linear system equivalent to our hypothesized color cube with $K = 128$, rather than $K = 5$, for calculation of gun-counts. Figure 5 shows a flow diagram for the steps by which the system reproduces the relative energy distribution of the original scene. Each step is postulated to be a linear transformation from MSS counts to gun-counts to film exposure to film transmission equivalent to original relative energy.

The hue distribution within cubic systems of the PFC type is shown in figure 6, (a) and (b). Figure 6(a) shows the cube standing on its black vertex, with the white vertex projecting toward the viewer. The other vertices of the cube are labeled with the primary hue which each represents, together with the relative energy quantities required to produce them. Thus, each point in the cube is the vector addition of the three relative energies and colors are combined by vector addition. Figure 6(b) is a side view of the cube which shows that the three additive primaries (blue, green, and red) are one-third of the way from black to white, whereas the three subtractive primaries (yellow, cyan, and magenta) are one-third of the way from white to black. This is logical since the additive primaries combine to form white, while the subtractive primaries combine to form black. The symmetry is completed by the fact that all complementary pairs are exactly opposite each other, e.g., yellow and blue, magenta and green, and cyan and red.

The same color distribution described above was presented in 1952 by Hickethier, based on empirical evidence in the printing industry, which utilizes subtractive primaries (Hickethier, 1952). Of course, the PFC is equivalent to a combination of additive primaries (gun-counts) and subtractive primaries (film emulsion dyes).

Color is usually considered to consist of three components, hue, value, and chroma. Hue represents the dominant wavelength, value is relative brightness, and chroma is the saturation or purity, vis-a-vis gray. The relationship between these components and the color cube are illustrated in figure 7. We have adopted the following conventions for mathematical designation of these properties. Hue is expressed in degrees (and letters) following the hue diagram shown in figure 8.



$$GC_{\text{pixel}} = \frac{128}{MSS} MSS_{\text{pixel}}$$

$$RE_{\text{pixel}} = \frac{5}{MSS} MSS_{\text{pixel}}$$

$$\overline{GC} = 128$$

$$GC_{\text{MAX}} = 255$$

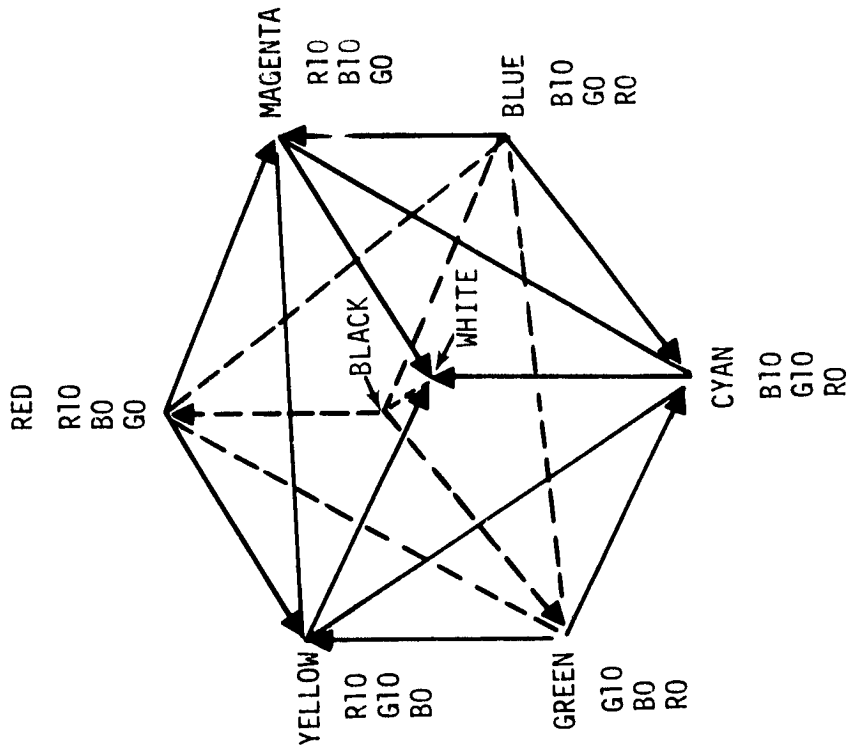
$$\overline{RE} = 5$$

$$RE_{\text{MAX}} = 10$$

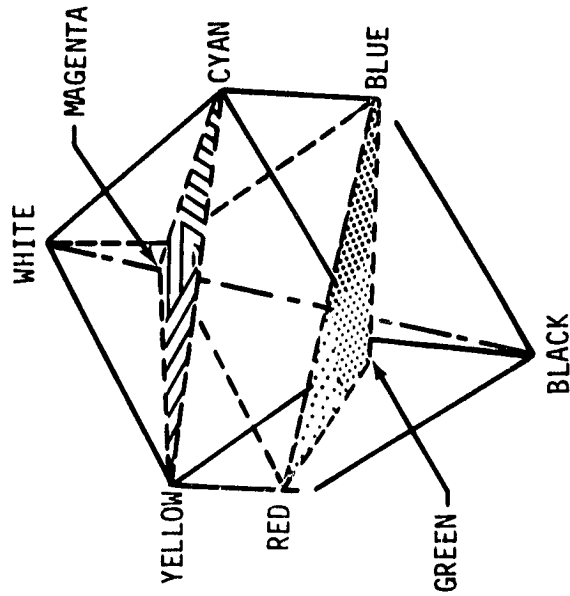
*Sensor counts expressed as multiples of scene mean.

Figure 5.- Color normalization and reproduction using relative energy (RE) concept.

LEGEND
 B = Blue
 G = Green
 R = Red

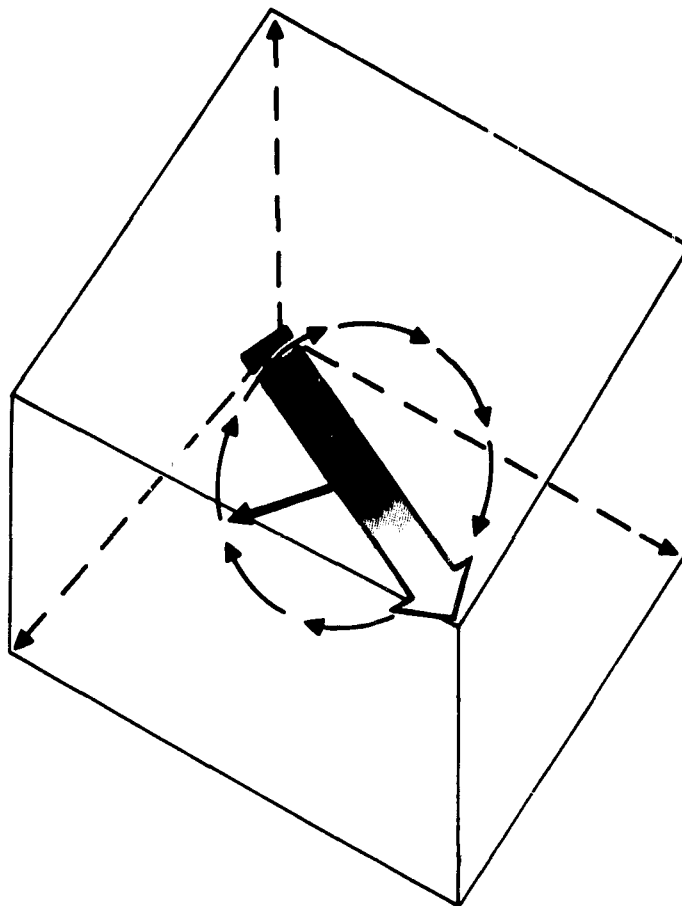


(a) Color cube with white vertex depicted toward viewer



(b) Side view of color cube resting on black vertex

Figure 6.— Hue distribution in the color cube.



- → → Hue Dominant Wave
- Chroma Grey → Saturated Color
- █ → Value Black → White

Figure 7.— Relationship of hue, value, and chroma notation to the color cube.

~~3-5~~
15

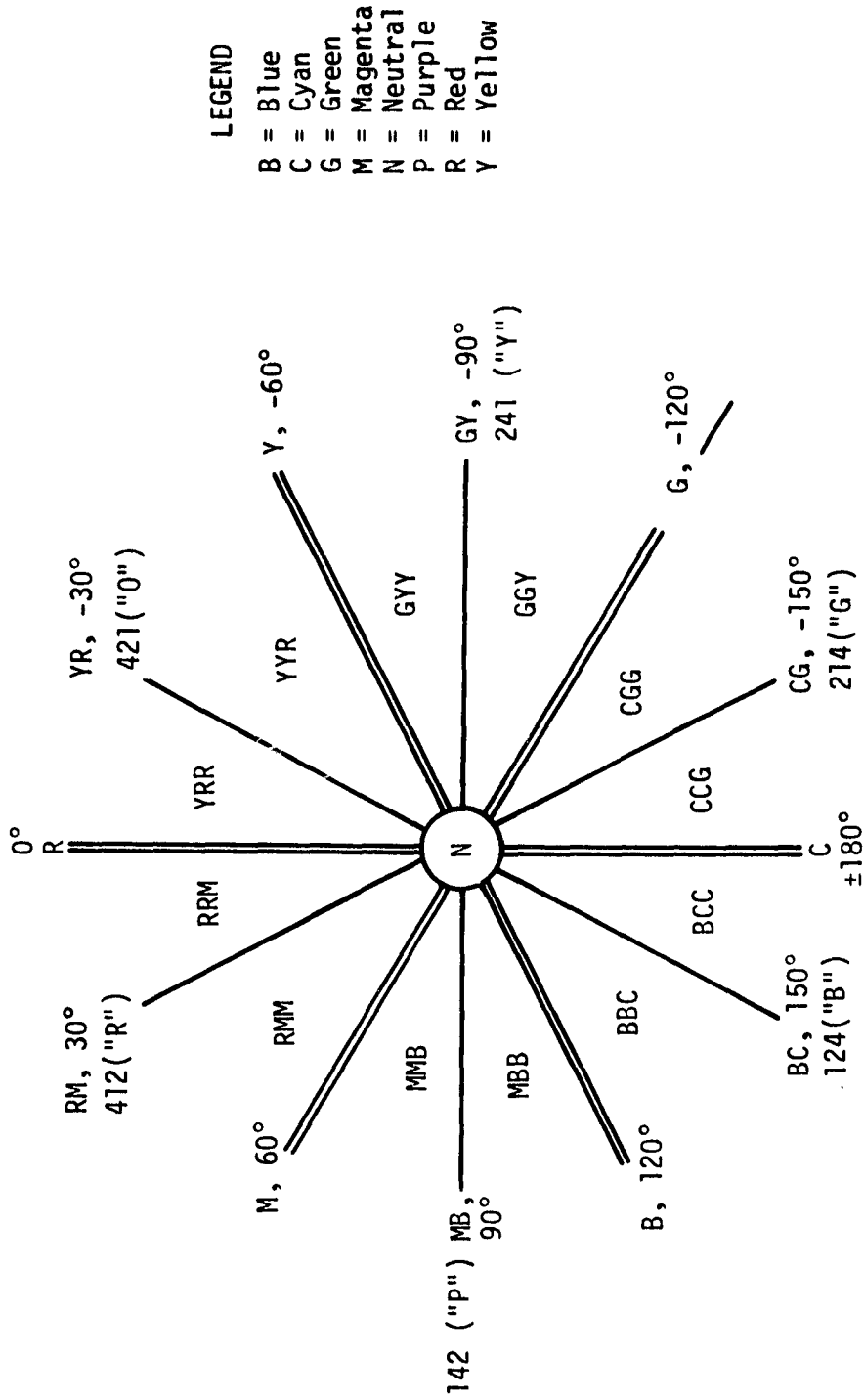


Figure 8.— Diagram of color cube hue notation.

Chroma is expressed in units of relative energy. (The maximum chroma possible is 8.165 and occurs at the primary apices of the cube.) Value is expressed on a scale of 0 to 10 along the black-white line, and each unit has a length of $\sqrt{3}$ relative energy units. The following are calculations for each of these color components.

$$\text{Value} = S/3$$

$$\text{Chroma} = \sqrt{A - 1/3S^2}$$

$$\text{Hue} = \cos^{-1} \left(\frac{2X_3 - X_1 - X_2}{2 \sqrt{A - X_1X_2 - X_1X_3 - X_2X_3}} \right)$$

where

$$A = X_1^2 + X_2^2 + X_3^2$$

$$S = X_1 + X_2 + X_3$$

$$X_1 = \text{relative energy for channel 1}$$

$$X_2 = \text{relative energy for channel 2}$$

$$X_3 = \text{relative energy for channel 4}$$

For interpretive purposes, we have found that it is helpful to segment the color cube according to channel ratios or rankings, since these may represent major differences which are directly attributable to physical causes such as the presence or absence of chlorophyll. The relationship between channel rankings and hues is shown in figure 9. These rankings have led us to hypothesize a generalized color sequence for agricultural crops. This is presented in table 3. Variations from this sequence, together with value and chroma differences, may serve to identify additional phenological changes and help to differentiate among crops. Once the species and growth stage of a crop have been determined, we think that residual color differences may reveal stress and damage, and that field texture may indicate stand deficiencies.

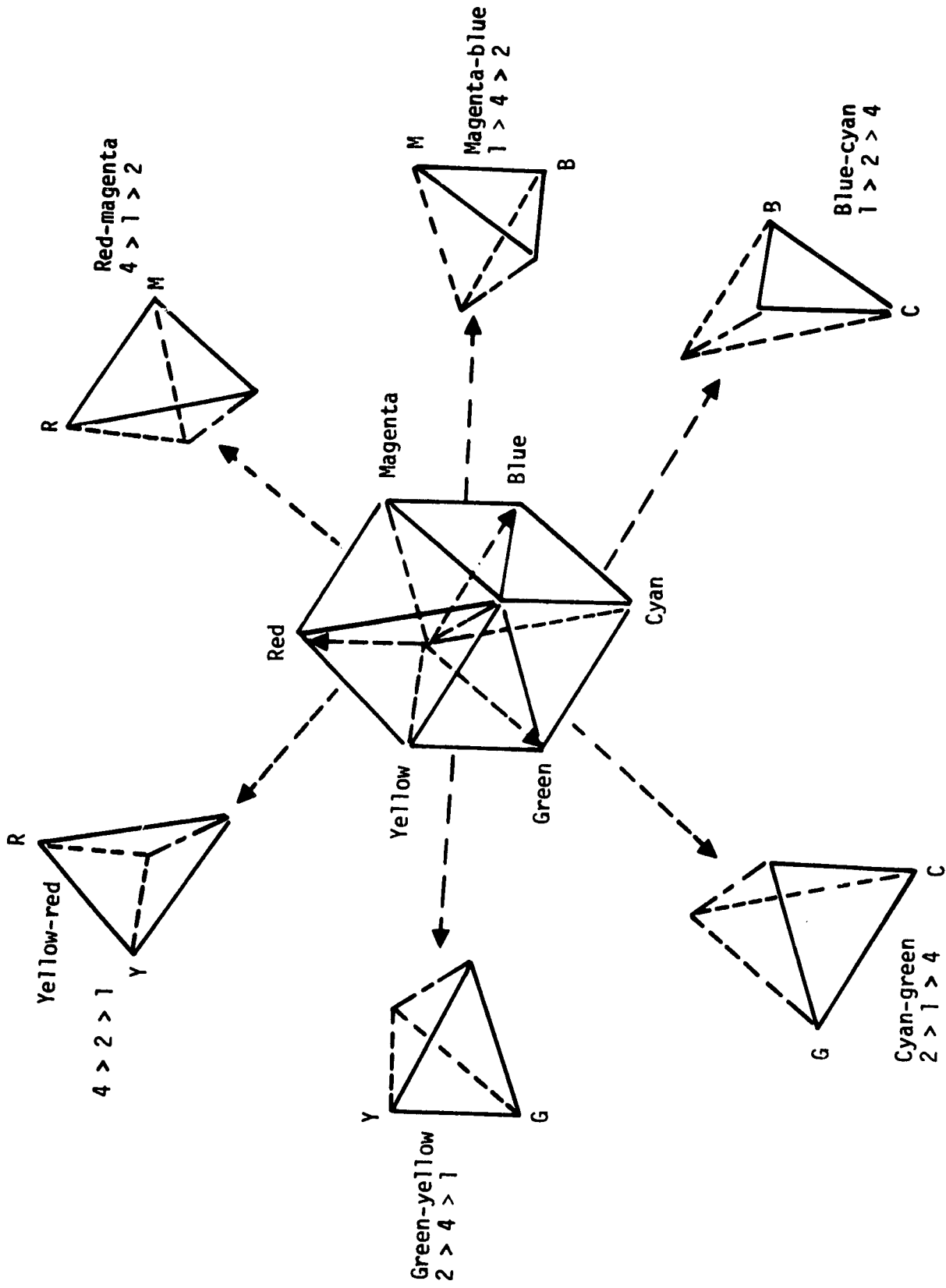


Figure 9.- Channel ranking hue domains of the color cube.

TABLE 3.- GENERALIZED COLOR SEQUENCE FOR AGRICULTURAL CROPS

<u>Crop description</u>	<u>Color components</u>	<u>Normalized rankings</u>	<u>Line printer pixel code*</u>	<u>Mnemonic</u>	<u>Hue in degrees</u>
Preplowing	Cyan-green	214	G	Green	-120 to -180
Postplowing	Blue-cyan	124	B	Blue	120 to 180
Increasing ground cover	Magenta-blue	142	P	Purple	60 to 120
Good ground cover	Red-magenta	412	R	Red	0 to 60
Reduction in chlorophyll	Yellow-red	421	O	Orange	0 to -60
Ripe	Green-yellow	241	Y	Yellow	-60 to -120
Bare soil/stubble	Cyan-green	214	G	Green	-120 to -180

*Shown in quotation marks in Figure 8.

The first large-scale test of the color sequence approach utilized spectral data and associated ground truth for 456 spring wheat fields distributed in 1978 over 35 sites in North Dakota, South Dakota, Montana, and Minnesota. Approximately 90 percent of these fields were found to follow our postulated color sequence during the normal spring small grain growing season. The remaining fields were unidentifiable due to insufficient spectral data as a result of clouds, etc. These results demonstrate that the proposed standardization and color transformation make possible the comparison of Landsat data acquired over widely varying points in space and time.

4. SUMMARY AND CONCLUSIONS

A method has been found for removing most of the scenewide acquisition-to-acquisition variability in Landsat data. The resulting mean-corrected values, called relative energies, can be further transformed into a color space which permits direct comparison of visual and computerized interpretations of MSS digital data. Preliminary tests indicate that the procedures discussed in this paper permit multitemporal normalization and classification of Landsat signatures over wide areas. Multitemporal color sequences derived by these techniques appear to correspond with ground-truth observations on crop development in the major spring wheat production area of the United States.

5. BIBLIOGRAPHY

Hickethier, A.: Color Matching and Mixing. (English translation.) Batsford, London, 1952.

Lambeck, P. F.; and Potter, J. F.: Compensation for Atmospheric Effects in Landsat Data. Proceedings of the LACIE Symposium, JSC-16015, 1979, pp. 723-738.

Land, H.: The Retinex Theory of Color Vision. Proc. Royal Inst. of Great Britain, vol. 47, 1974, pp. 23-58.

Land, H.: The Retinex Theory of Color Vision. Sci. American, vol. 237, 1977, pp. 108-128.

Munsell: Book of Color. Macbeth Div. of Kollmorgen Corp., 2441 N. Calvert Street, Baltimore, Maryland 21218, 1963.

APPENDIX A

CORRELATION STRUCTURE IN THE NORMALIZED DATA

CORRELATION STRUCTURE IN THE NORMALIZED DATA

Let

$$X_{ijk} = \begin{bmatrix} X_{ijk1} \\ X_{ijk2} \\ \vdots \\ X_{ijkp} \end{bmatrix}$$

be the measurement vector for k th pixel of j th field in i th scene. Define

$$Y_{ijk} = \frac{X_{ijk}}{\bar{X}_{i..}} = \begin{bmatrix} \frac{X_{ijk1}}{\bar{X}_{i..1}} \\ \frac{X_{ijk2}}{\bar{X}_{i..2}} \\ \vdots \\ \frac{X_{ijkp}}{\bar{X}_{i..p}} \end{bmatrix}$$

where $\bar{X}_{i..}$ is the scene mean. Due to a large number of pixels in a scene

$$\bar{X}_{i..} \doteq \mu_i = \begin{bmatrix} \mu_{i1} \\ \mu_{i2} \\ \vdots \\ \mu_{ip} \end{bmatrix}$$

where μ_i is the actual mean of the population of pixels represented by the i th scene. (Dot on top of equality sign means approximation.)

Let n_{ij} be the number of pixels in j th field of scene i . Denote

$$\bar{X}_{ij.} = \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} X_{ijk}$$

$$\bar{X}_{i..} = \frac{1}{\sum_j n_{ij}} \sum_j \sum_k X_{ijk}$$

$$\bar{X}_{...} = \frac{1}{\sum_i \sum_j n_{ij}} \sum_i \sum_j \sum_k X_{ijk}$$

$$Y_{ij.} = \frac{5\bar{X}_{ij.}}{\mu_i}$$

$$Y_{i..} = \frac{5\bar{X}_{i..}}{\mu_i} \doteq 5$$

$$Y_{...} = 5 \frac{\bar{X}_{i..}}{\mu_i} \doteq 5$$

For the transformed data, the estimated covariance matrix is given by

$$\begin{aligned} S_{\ell,m} &= \frac{1}{\sum_{i,j} n_{ij} - 1} \sum_i \sum_j \sum_k \left[(Y_{ijk\ell} - Y_{... \ell}) (Y_{ijkm} - Y_{... m}) \right] \\ &= \frac{25}{\sum_i \sum_j n_{ij} - 1} \sum_i \sum_j \sum_k \left(\frac{X_{ijk\ell}}{\mu_{i\ell}} - 1 \right) \left(\frac{X_{ijkm}}{\mu_{im}} - 1 \right) \\ &= \frac{25}{\sum_i \sum_j n_{ij} - 1} \sum_i \frac{1}{\mu_{i\ell} \mu_{im}} \sum_j \sum_k \left[(X_{ijk\ell} - \mu_{i\ell}) (X_{ijkm} - \mu_{im}) \right] \\ &= \frac{25}{\sum_i \sum_j n_{ij} - 1} \sum_i \frac{\sum_j n_{ij} - 1}{\mu_{i\ell} \mu_{im}} S_{\ell,m}^{(i)} \end{aligned}$$

Where

$$S_{\ell,m}^{(i)} = \frac{1}{\sum_j n_{ij} - 1} \sum_j \sum_k (X_{ijk\ell} - \mu_{i\ell}) (X_{ijkm} - \mu_{im})$$

which is the within-segment covariance matrix for scene i . Thus,

$$S_{\ell,m} = \frac{25}{\sum_i \sum_j n_{ij} - 1} \sum_i \sum_j (n_{ij} - 1) \frac{S_{\ell,m}^{(i)}}{\mu_{i\ell} \mu_{im}}$$

The correlation ($\rho_{\ell,m}$) between channels ℓ and m given by

$$\rho_{\ell,m} = \frac{S_{\ell,m}}{\sqrt{S_{\ell,\ell} S_{m,m}}}$$

does not depend upon the mean vector μ_i . It follows that

$$\rho_{\ell,m} = \frac{\sum_i \sum_j (n_{ij} - 1) S_{\ell,m}^{(i)}}{\sqrt{\left[\sum_i \sum_j (n_{ij} - 1) S_{\ell,\ell}^{(i)} \right] \left[\sum_i \sum_j (n_{ij} - 1) S_{m,m}^{(i)} \right]}}$$

Accordingly, the correlation structure is not changed at all by the transformation of X_{ijk} to Y_{ijk} .