On Displaying Multispectral Imagery

Jack Bryant
Center for Approximation Theory, Department of Mathematics, Texas A&M University, College Station, TX 77843

ABSTRACT: A digital image with more than three measurements per picture element is not readily displayed on a conventional color monitor. Understanding of the image is hampered by the lack of a single image product. A method is proposed to solve these problems. A transformation from the multispectral image to a displayable three band image is developed which preserves the measurement-space structure of the data. Color selections are stable and reproducible: modest changes in the sensor or scene cause inperceptible changes in the displayed colors. Image processing operations such as clustering and classification can be performed on the three-band image instead of the original higher dimensional data with essentially identical results. The entire process is unsupervised.

INTRODUCTION

LET US BEGIN with a large multi-dimensional digital image such as one produced by a remotely positioned multispectral sensor. Our ultimate concern is to understand the scene remotely sensed from the digital image. Human intervention is certain to be required. The only avenue available for presenting the picture as a whole to a human analyst is visual. Suppose the display device is a color (RBG) monitor with 512 by 512 picture elements (pixels). Each displayable pixel is represented by three eight-bit bytes correspond the three colors red, green, and blue. The display device is connected to a general purpose image analysis system. Although the methods proposed here apply to display systems with different capabilities, it serves our ends to examine the problem in this setting. With this in mind, consider the problem of presenting the human visual system with an image which conveys enough of the relationships in the scene to support understanding.

Several formidable problems arise. The multispectral sensor produces more bands than can be displayed at one time. Thus, one must select or otherwise combine in some way the original bands to produce a displayable image. Even the simplest alternative (the selection of three) is complex: there are 210 ways of selecting three bands from seven and assigning the three display channels to the bands selected (Sheffeld, 1985). In the example presented later, there are 11 bands, giving 990 potential assignments. The temptation to produce exotic displays is irrestible; the number of ways of sampling, rotating, selecting, transforming, and clustering is overwhelming. A general image analysis system must furnish these options, but it is wasteful to apply them exhaustively.

In addition, many of the analysis tools are not available if the image is spatially larger than the 512 by 512 display area. For example, an automatic scaling operation applied to subsets of an image will lead to different "colors" for the same object indifferent segments. The transformation that sales is "trained" on the particular subset instead of on the whole image. Different subsets of large images are certain to contain different distributions of measurements. It may be *possible* to use the same scaling procedure on the entire image, but this feature is not usually *built in* and, therefore, it is unlikely to be used by an application-oriented analyst.

Another problem is the quality of data: while the precision of each display channel is eight bits, true eight-bit multispectral measurements are rare. Scaling or histogram equalization may enhance the appearance of the display, but one wonders what such operations do to the data. Here again, there is a "mismatch" between the display device and the data. The display needs higher precision data in fewer bands than real data.

To one who must interpret the multispectral measurements, an almost utopian display product would take the raw data and transform it into a single image which aids understanding, rather than a large set of possibly helpful products. It would be a definite advantage if the "data structure" were preserved in the displayable image. (Unfortunately, there is no clear understanding extant of the *meaning* of the term "structure." I shall propose one later in this paper, but until then I use the term informally.) To be useful, the colors assigned in the display should convey some of the meaning of the original, and should be "reproducible" – that is, they should not change much, given modest changes in the sensor or scene.

The problem we undertake thus has three related aspects. We are given a digital multi-image containing $N>3$ bands. We seek

- to reduce the original dimensionality from N to three while preserving the structure in N dimensions;
- to make a displayable color product with reproducible colors; and,
- to effect these ends with an essentially automatic (i.e., unsupervised) computer program.

FEATURE SELECTION

Feature selection is a technique for reducing the dimensionality of essentially arbitrary multi-dimensional data. A widely-used method is the principal components transformation (Taylor, 1974; Merembeck *et al.*, 1977; Lillesand and Kiefer, 1979, for example). The reduced dimensionality data presumably preserve the "structure" of the original data. The problem with applying this technique to the image display problem is simply stated: how does one assign RBG channels to the reduced three-band image? Another problem is: how does one scale the transformed image?

The simple method of scaling each transformed band to fill the displayable range (i.e., the set $\{0, 1, \ldots, 255\}$ of numbers the display system understands in each of the three colors) produces disappointing display products. The colors are not predictable; small changes in the input scene lead to wild changes in in the output colors. There remain six ways to choose the three color assignments to the three reduced dimensions. This undesirable behavior so hampers image understanding that the original multispectral measurements, present as several distinct images, must be constantly at hand in order to interpret the reduced one. Thus, no resources have been saved: rather than fewer images to understand, the user has more. This is pointed out by Hay and Thomas (1977).

The objective in the principal components approach is to select an orthogonal set $\{x_1, x_2, x_3\}$ of three vectors which minimizes mean-square truncation error in the expansion of the prototypes in an orthogonal series beginning with these three vectors. While some measure of accuracy is provided, the measure is not as directly tied to the structure question as one might desire. (By examining the remaining $N - 3$ eigenvalues of the covariance

PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING,
Vol. 54, No. 12, December 1988, pp. 1739–1743

0099-1112/88/5412–1739$02.25/0
©1988 American Society for Photogrammetry and Remote Sensing

matrix of the prototypes, one obtains an estimate of how much "variance" is modeled by the initial terms of the series.) However, one vital property is present: the transformation is easily and quickly applied, not only to the prototypes, but to *all* data, as a $3 \times N$ matrix multiplication. That is, the dimensionality reduction step is a linear transformation from R^N to R^3. In spite of this, the method considered below, which produces displays, performs as well on the transformation given by the principal components mapping. A slight improvement can be obtained, and this sheds some light on one possible meaning of "structure" as applied to image analysis problems.

One way to look at structure is from the point of view of a *classifier*. Imagine a classifier in N dimensional space, and consider the classification of the prototypes. Suppose that, after the prototypes were transformed to the lower dimensional space (here of dimension three), the classification of the three-dimensional transformed prototypes remained the same as the original. (This requires that the classifier also be somehow transformed to the lower dimensional space.) In that case, one might say the original set of prototypes *was* three-dimensional, at least from the point of view of that particular classifier.

For example, if the classifier is a nearest-neighbor classifier, then the classifier is determined by the attractors (sometimes casually called class means or centers) and the distance function. The author believes that, in the absence of "expert" knowledge, the correct distance function is Euclidean distance. The three-dimensional class attractors are easily obtained using the same transformation which produced the transformed prototypes, and Euclidean distance in R^3 can serve as the distance function. Thus, the classifier in R^3 is defined. It is probable that a transformation which preserves distances between prototype pairs in N-dimensional space in the reduced three-dimensional space will not change any classification, no matter what the attractors, and this is assumed if the class attractors are themselves prototypes. Let us make the idea suggested here more precise.

A linear transformation from R^N to R^3 is a $3 \times N$ matrix $\mathbf{L}$. From the set of n distinct prototypes form the $p = n(n-1)/2$ pair differences $\mathbf{z}_k = \mathbf{P}_i - \mathbf{P}_j$, $1 \le i < j \le n$. The measure of how well a matrix $\mathbf{L}$ preserves the distances between the prototypes is the function F of $3N$ variables, defined by

$$F(\mathbf{L}) = \sum_{k=1}^{p} (\|\mathbf{z}_k\| - \|\mathbf{L}\mathbf{z}_k\|)^2.$$

In this, $\|\mathbf{z}_k\|$ is the Euclidean distance in R^N between prototypes and $\|\mathbf{L}\mathbf{z}_k\|$ is the distance in R^3 between transformed pairs. This *objective function* has been considered before (Bryant and Guseman, 1979). The minimization of F is obtained by an iterative steepest descent method, modified to account for the unfortunate fact that F is not differentiable everywhere.* The starting point for the iterative method is the principal components map trained on the prototype pair differences.

If we can confidently assume the set of prototypes represent the measurement space structure of the image, then a linear mapping from R^N to R^3 which preserves the distances between prototypes will automatically preserve the structure of the N-band image in R^3. Such a mapping will be a congruence relation from R^N to R^3 when restricted to the set of prototypes. In practice, the mapping will not exactly be a congruence relation, but it will be as close as possible to one, and will almost certainly not change any classification of a prototype. Moreover, because of linearity, points other than the prototypes are easily mapped to R^3. The problem, then, is to choose representatives of the classes present in the image.

*However, F is differentiable at a local extremum.

PROTOTYPES

The first step in the solution of the problem is the automatic estimation of boundary image elements (pixels) in the N band image. It is essential that spatial boundaries be avoided when sampling an image to obtain representatives of the classes present, for the measurements obtained by the sensor do not belong to any actual object. As elements of R^N, boundary pixels are distinguished by being significantly different in at least one band from one of their spatial neighbors. The thresholds (which serve to define "significant") are initially set so the image is mostly boundary, but adapt rapidly to the actual variability of the data. Be default, 45 percent of the pixels are sought to be boundary points, so that the boundary is not expected to be small or "thin."

After the boundary is identified, what remains is a collection of blobs separated by boundary. If the resolution of the sensor is adequate to sample the actual objects in the scene, then one must assume that the spectrally homogeneous blobs represent the data.

From the blobs it is relatively easy to extract n samples so that the samples are distinct and each blob is represented in one sample. Call the n samples *prototypes;* they form a set $\{\mathbf{P}_1,\ldots,\mathbf{P}_n\}$ of n elements of N-dimensional Euclidean space R^N. In the specific application under discussion, they are selected in the course of clustering the original N-band image. The program used, Version 12 of AMOEBA (Bryant, 1979; Jenson *et al.*, 1982), begins by locating spatially connected regions with relatively homogeneous mesurement-space behavior. Samples from the regions are clustered by a method which begins with many attractors and reduces the number while attempting to fit a spatial-spectral model for the data. When 32 attractors remain, the process is interrupted and the color display module is entered.

COLOR DISPLAYS

The transformation $\mathbf{L}$ takes the original measurements $\mathbf{p}$ into 3-vectors $\mathbf{q} = \mathbf{L}\mathbf{p}$. We wish to display the transformed image. Recall the display device has three 8-bit channels. If we denote the finite set $\{0,1,\ldots,255\}$ by D, then we actually seek a mapping from the set containing the data into D^3. As might be expected, simply translating and scaling the transformed measurements $\mathbf{q}$ is not sufficient. The objections raised in the introduction apply here: the color display must be arranged with more care. Note also that scaling is sure to destroy the relative distances between prototypes unless the same scale factor is applied to each of the three transformed bands.

We seek, then, attributes of the original image which can be found automatically to define the R^3 to D^3 transformation. In this search, let us keep in mind the main application, to multispectral measurements organized as images. A pixel consists of N measurements of reflectance or radiation from the object scene obtained over different parts of the electromagnetic spectrum. Presently, essentially all such measurements are at most eight bits in precision. That is, the original pixels $\mathbf{p}$ are contained in D^N. Through the n prototypes $\mathbf{p}_i$ we known exactly where in D_N the data lie, and can reduce the dimensionality linearly to R^3 preserving distances between prototypes with small errors. Our search for attributes to preserve is guided by a desire to make the displayed colors natural, in the sense that the image produced is similar to a familiar product. Color infra-red (IR) film is one such product.

To this end, locate the prototype $\mathbf{P}_w$ with smallest norm in R^N – that is, which minimizes

$$\|\mathbf{p}\| = \sqrt{\frac{1}{N}\sum_{j=1}^{N} p_j^2}$$

(here $\mathbf{p} = (p_1,\ldots,p_N)^T$). (Water(if present) or wet land will usually be dark, but perhaps not so dark as shadows.) Define the transformation $\mathbf{M}$ by $\mathbf{M}(\mathbf{p}) = \mathbf{L}(\mathbf{p}-\mathbf{p}_w)$; $\mathbf{M}$ is not linear but it preserves distances between prototypes exactly like $\mathbf{L}$ does. Note that $\mathbf{M}(\mathbf{p}_w) = 0$.

Next, locate the prototype $\mathbf{P}_u$ with maximum R^N norm. (If an urban class or bare soil is present, it will often appear bright, but perhaps not as bright as a cloud or the sun's reflection from a body of water if either is present.) In each case which has been studied, the prototypes $\mathbf{p}_w$ and $\mathbf{p}_u$ are extreme points of the convex hull of the set of prototypes. In most cases, they are points most separated—that is, the distance between these two prototypes is the diameter of the set of prototypes.

Let $\mathbf{q} = \mathbf{M}(\mathbf{p}_u) \epsilon R^3$. Let $\mathbf{R}_1 : R_1 \rightarrow R^3$ be a rotation such that $\mathbf{R}_1$ rotates $\mathbf{q}$ so that the three image coordinates are equal and positive. (The construction of such a rotation is given in the appendix.) In all cases which have been studied, the image under the product $\mathbf{R}_1\mathbf{M}$ of the set of prototypes lies in the first octant—that is, the subset of three-dimensional space with all coordinates non-negative. Thus, the transformed prototypes can be displayed without clipping by scaling. The selection of "colors" is certainly not clear, and slight changes in the image make large changes in any arbitrary selection of RGB assignments to be transformed-rotated-scaled measurements. We now describe how the colors are assigned.

Many users of multispectral imagery will possess at least one band (say band v) in the visible part of the spectrum and another (band r) in the near IR. For many years these data have been displayed by assigning the red display channel to a near IR band and the green and blue to visible bands.* The resulting display looks like color IR film; healthy vegetation will appear red. The ratio of the IR band to a visible band is a good indicator of vegetation. A better indicator for scenes which may contain urban classes can be obtained by the following procedure. A scaled ratio of the IR measurement to the sum of visible measurement and the norm of the prototype is an improved indicator. Select the prototype $\mathbf{p}_v$ different from $\mathbf{p}_w$ and $\mathbf{p}_u$ in which this ratio

$$\frac{p_r}{2p_v + \|\mathbf{p}\|}$$

is largest. (To be sure, the user must supply band numbers of a visible band and a near IR band, but the process in otherwise unsupervised.)

The point $\mathbf{r} = \mathbf{R}_1\mathbf{M}(\mathbf{p}_v)$ almost certainly lies in the first octant in R^3. Let $\mathbf{R}_2$ be the rotation about $(1,1,1)^T$ which maps $\mathbf{r}$ into the plane determined by the two vectors $(1,0,0)^T$ and $(1,1,1)^T$ which maximizes the first (red) coordinate of $\mathbf{R}_2\mathbf{r}$. (Details on how the rotation can be constructed are given in the Appendix.) The composition $\mathbf{T} = \mathbf{R}_2\mathbf{R}_1\mathbf{M}$ takes the set of prototypes into the first octant in R^3. To the extent that $\mathbf{L}$ preserved distances between prototypes, the affine mapping $\mathbf{T}$ also preserve distances.

A scaling operation is now applied to bring the transformed prototypes into D^3. Determine the transformed measurements $\mathbf{t}_i = \mathbf{T}\mathbf{p}_i$. Each $\mathbf{t}_i$ is a 3-vector. Find the three smallest intervals (l_c, m_c), $c = 1,2,3$, which contain the components of the $\mathbf{t}_i$. For example, l_1 is the minimum transformed red coordinate and m_1 is the maximum. Let $d = \max_{c=1,2,3}\{m_c - 1_c\}$. Let $\alpha = 200/d$. Define a bias vector $\mathbf{b} = (30,30,30)^T$. Under the transformation $\mathbf{U}(\mathbf{p}) = \alpha\, \mathbf{T}(\mathbf{p}) + \mathbf{b}$, relative interprototype distances are preserved, and the set of prototypes maps comfortably inside the display space D^3. Experience has shown essentially no clipping (or folding) is required in the application of this map to all of the data.

*Some applications will make different color assignments; in that case, band r should be the band displayed red and band v the band displayed green. Then the process outlined here will produce a composite display similar to the one expected which contains all the class separation structure of the original.

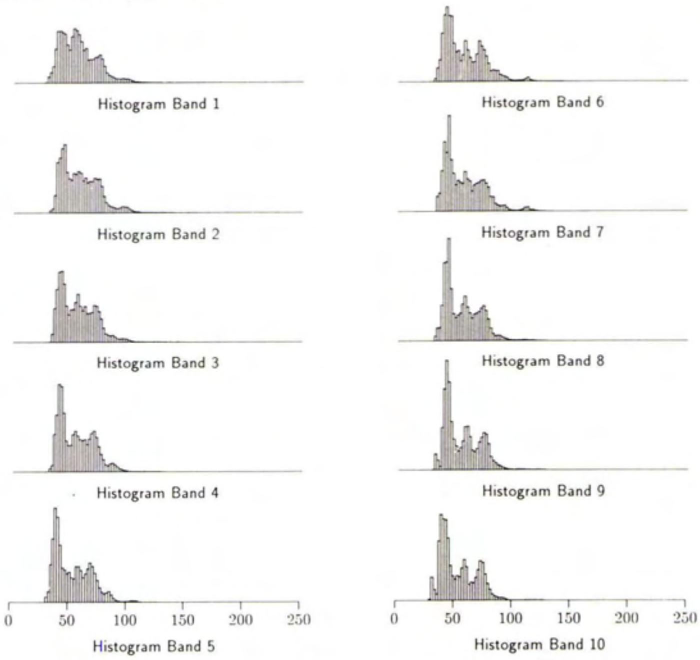


FIG. 1. (a) Histograms, bands 1 to 5, input data. (b) Histograms, bands 6 to 10, input data.

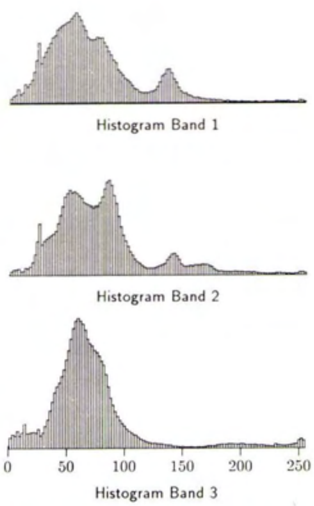
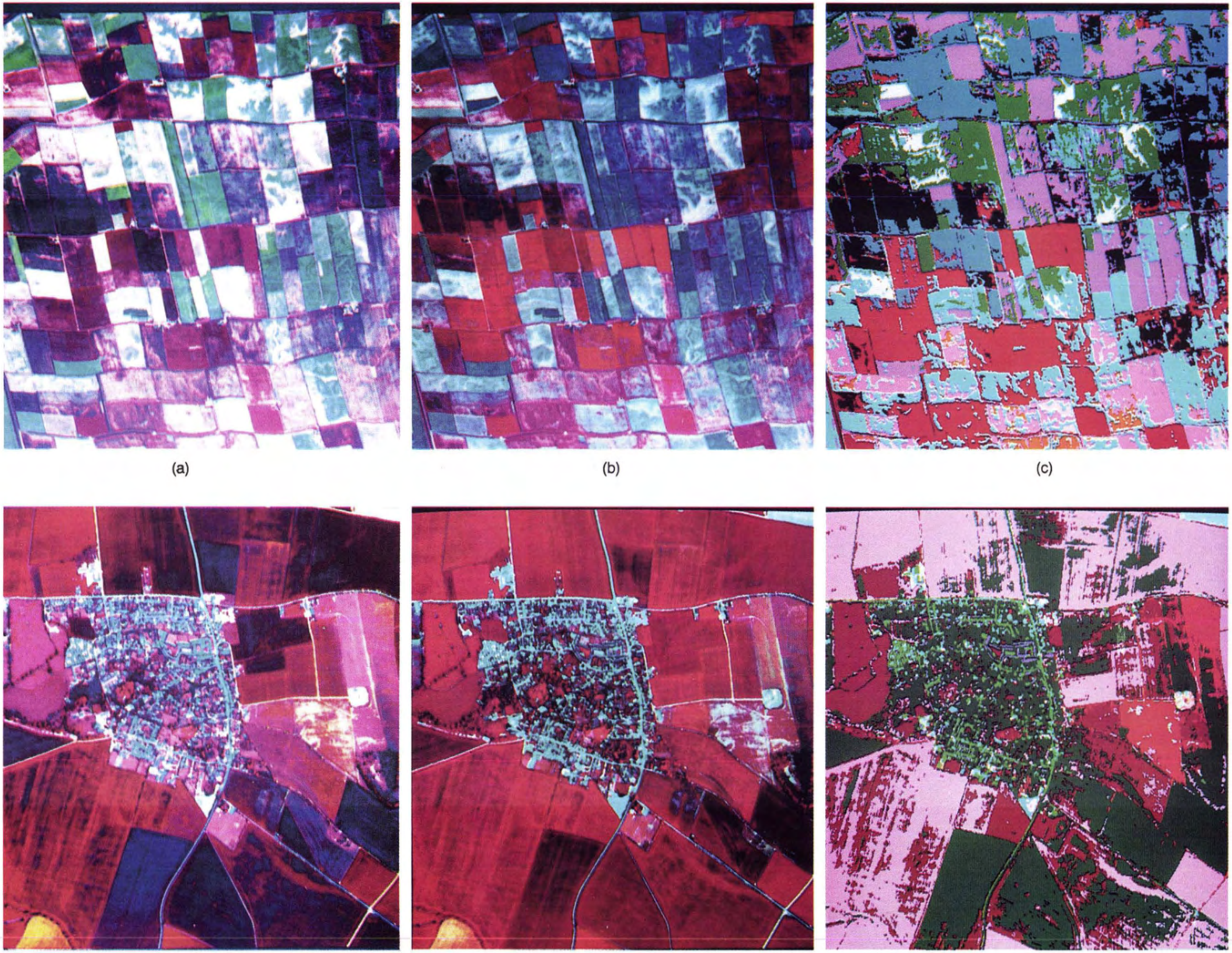


FIG. 2. Histograms, reduced dimensionality data.

EXAMPLE APPLICATIONS

Two typical examples are presented here in Plate 1. The eleven-band 512 by 512 images were lent to the author by the Natural Environmen

tal Research Council Unit for Thematic Information, Department of Geography, Whiteknights (Reading, U.K.), John Townshend, Director. The first image displayed is bands 6, 4, and 3 displayed RGB. This image is scaled to fill the range 0 to 255 of the I² S Model 70. The next is the unscaled transformed image prepared by the method described here; visible band index 3



(a) (b) (c)

(d) (e) (f)

PLATE 1. (a) Bands 6, 4, 3 displayed RGB, scaled. (b) Composite of 11 bands, not scaled. (c) False color cluster map. (d) Bands 6, 4, 3 displayed RGB, scaled. (e) Composite of 11 bands, not scaled, (f) False color cluster map.

and IR band index 6 were used to position the reduced dimensional image in display space. The third image is the cluster map output from AMOEBA; the colors have no significance. Each image was photographed directly from the display screen.

Also of interest are the one-dimensional histograms obtained. For two of the images, the one-dimensional histograms of typical input bands and of the display product are shown. In Figure 1, typical histograms of bands 1 to 10 are shown. Figure 2 shows the three bands of a typical output product.

CONCLUSIONS AND DIRECTIONS OF FUTURE WORK

A digital image with more than three measurements per picture element is not readily displayed on a conventional color monitor; it is difficult to interpret because only three bands can be presented at once as an image, so the analyst must shuffle between several display products. We have introduced a transformation from the multispectral image to a displayable three-band image which preserves the measurement-space structure of the data in the sense that relative distances between automatically selected prototypes are preserved in the three-band false color image. Color selections are stable and reproducible in the sense that small changes in the sensor or scene cause inperceptible changes in the displayed colors. Image processing operations such as clustering and classification can be performed on the three-band image instead of the original higher dimensional data with essentially identical results. The entire process is unsupervised.

The method has been tested on several multispectral scanner data sets, and preliminary results were presented. The saving in processing time on the three-rather than eleven-band test images more than repays the time taken to find and apply the dimensionality reduction transformation. Research is now being undertaken on the question of reconstructing approximately the original image form the reduced displayable image. It appears that this should be possible with subjectively negligible errors. If this proves to be the case, then new light will be shed on the meaning of the term "structure."

ACKNOWLEDGMENTS

I was fortunate to have the cooperation of Professor Townshend's group of NERC; in particular, the assistance of Mr. J. J. Settle is gratefully noted. The Department of Electrical Engineering at Texas A&M allowed me access to their I²S-VAX 11/750 image display system. Two Reviewers of the draft made suggestions which helped me claifry the presentation. It is a pleasure to acknowledge this support.

REFERENCES

Bryant, Jack, 1979. On the clustering of multidimensional pictoral data, *Pattern Recognition* 11:115–125.

Bryant, Jack, and L. F. Guseman, Jr., 1979. Distance preserving linear feature selection, *Pattern Recognition* 11:347–352.

Jenson, Susan K., Thomas R. Loveland, and Jack Bryant, 1982. Evaluation of AMOEBA: a spectral-spatial classification method, *J. Applied Photographic Engineering* 8:159–162.

Hay, C. M., and G. S. Thomas, 1977. *Development of Techniques for Producing Strata Maps and Development Photointerpretive Methods Based on Multitemporal Landsat Data,* Final Report for NASA Contract NAS9-14565, Remote Sensing Research Program, Berkeley, California.

Lillesand, T. M., and R. W. Kiefer, 1979. *Remote Sensing and Image Interpretation,* Chapter 10; John Wiley and Sons.

Merembeck, B. F., F. Y. Borden, M. H. Podwysocki, and D. N. Applegate, 1977. Application of canonical analysis to multispectral scanner data, *Proceedings of the 14th Annual Symposium on Computer Applications in the Mineral Industries,* Society of Mining Engineers, American Institute in Mining, Metallurgical and Petroleum Engineers, New York.

Sheffield, Charles, 1985. Selecting band combinations from multispectral data, *Photogrammetric Engineering and Remote Sensing* 51(6):681–687.

Taylor, M. M., 1974. Principal components colour display of ERTS imagery; *Third Earth Resources Technology Satellite Symposium,* Section B, Paper I 12:1877–1897, Dec. 10–14, 1973, Washington, D. C.

(Received 3 March 1988; accepted 27 April 1988; revised 17 June 1988)

APPENDIX

Here we outline the construction of the two rotations mentioned above. The first rotation is essentially arbitrary, but the choice of the axis of rotation described now is a natural one. Denote by $\mathbf{x} \times \mathbf{y}$ the vector (cross) product and $\mathbf{x} \cdot \mathbf{y}$ the scalar (dot) product of vectors $\mathbf{x}$ and $\mathbf{y}$ in R^3. Recall that if $\mathbf{P}$ is orthogonal then $\mathbf{P}^{-1} = \mathbf{P}^T$.

Recall $\mathbf{q}$ is the transformed urban-like prototype. Let $\mathbf{d} = (1,1,1)^T/\sqrt{3}$, $\mathbf{b} = (\mathbf{q} \times \mathbf{d})/\|\mathbf{q} \times \mathbf{d}\|$, and $\mathbf{c} = \mathbf{d} \times \mathbf{b}$. The matrix $\mathbf{P}$ constructed with columns $(\mathbf{d},\mathbf{b},\mathbf{c})$ is orthogonal. Let $c = \mathbf{u} \cdot \mathbf{d}/\|\mathbf{u}\|$ and $s = \sqrt{1-c^2}$. It is easy to see that $c = \cos\theta$, $s = \sin\theta$, where θ is the angle to rotate about $\mathbf{b}$ so that the rotated $\mathbf{q}$ lies on $\mathbf{d}$. The matrix

$$\mathbf{R}_P = \begin{bmatrix} c & 0 & s \\ 0 & 1 & 0 \\ -s & 0 & c \end{bmatrix}$$

is a rotation about $\mathbf{b}$ in the $\mathbf{P}$ coordinate system which will make $\mathbf{u}$ point in the $\mathbf{d}$ direction. In the natural coordinate system, the rotation is $\mathbf{R}_1 = \mathbf{P}^{-1}\,\mathbf{R}_P\,\mathbf{P}$.

Now apply this rotation of the transformed vegetation-like prototype $\mathbf{v}$: $\mathbf{v}_r = \mathbf{R}_1\mathbf{v}$. Define a new coordinate system by the orthogonal matrix

$$\mathbf{Q} = \begin{bmatrix} 1/\sqrt{3} & 2/\sqrt{6} & 0 \\ 1/\sqrt{3} & -1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & -1/\sqrt{6} & -1/\sqrt{2} \end{bmatrix}.$$

Let $\mathbf{v}_e$ be the projection of $\mathbf{v}_r$ onto the plane determined by the last two columns of $\mathbf{Q}$. Let $c = \mathbf{v}_r \cdot \mathbf{b}/\|\mathbf{v}_e\|$ and $s = \mathbf{v}_r \cdot \mathbf{c}/\|\mathbf{v}_e\|$, where $\mathbf{b}$ is the second column of $\mathbf{Q}$ and $\mathbf{c}$ is the third. Note $c^2 + s^2 = 1$. Then

$$\mathbf{R}_Q = \begin{bmatrix} c & 0 & s \\ 0 & 1 & 0 \\ -s & 0 & c \end{bmatrix}$$

is a rotation in the $\mathbf{Q}$-coordinate system which places $\mathbf{v}_r$ in the plane determined by $\mathbf{d}$ and $\mathbf{b}$ nearest to $\mathbf{b}$; such a vector will have the second and third components equal and as large as possible (under rotations). Again, the natural coordinate system rotation is $\mathbf{R}_2 = \mathbf{Q}^{-1}\,\mathbf{R}_Q\mathbf{Q}$.

FOR SALE

Wild PUG IV Point Marker and Kern CPM-1 Comparator Point Marker

Address all inquiries c/o Box B, ASPRS, 210 Little Falls Street, Falls Church VA 22046