A Physically-Based Transformation of Thematic Mapper Data—The TM Tasseled Cap

ERIC P. CRIST AND RICHARD C. CICONE

Abstract—In an extension of previous simulation studies, a transformation of actual TM data in the six reflective bands is described which achieves three objectives: a fundamental view of TM data structures is presented, the vast majority of data variability is concentrated in a few (three) features, and the defined features can be directly associated with physical scene characteristics. The underlying TM data structures, based on three TM scenes as well as simulated data, is described, as are the general spectral characteristics of agricultural crops and other scene classes in the transformed data space.

I. INTRODUCTION

A NALYSIS of remotely sensed scanner data requires synthesis of information contained in several discrete spectral bands into information which can be associated with the physical characteristics of scene classes. This process might be thought of in three parts: 1) understanding the relationships among the spectral bands for the scene classes of interest, 2) compressing the n spectral bands of information into a manageable number of features, and 3) extracting physical scene characteristics from the spectral features. Application of Landsat-MSS data to vegetation monitoring (forests, agriculture, etc.) has included a variety of approaches to accomplishing one or more of the necessary steps in the process. Ratios of bands such as MSS7/MSS5 or the Normalized Difference (Rouse et al. [1]) are examples of features which partially fulfill these functions: they are based on known spectral interactions in green vegetation canopies, provide some limited data volume compression, and are more readily associated with some attributes than are the raw MSS bands. Principal components analysis provides data volume reduction, but presents substantial obstacles with regard to physical interpretation of the derived features, particularly between dates or scenes.

The Tasseled Cap transformation of MSS data, developed by Kauth and Thomas [2], accomplishes all three functions. Analysis of Landsat data from agricultural regions has shown that, on any given date, the four-band MSS data primarily occupy a single plane, with the various band pairs providing skewed views of that plane. This planar distribution of the data results from correlations between the two visible bands and the two infrared bands which arise as a result of plant and soil reflectance properties. The Tasseled Cap transformation rotates the MSS data plane such that the vast majority of data variability is concentrated in two features, i.e., the plane is viewed “head-on.” For a single plane, this view and one which presents the edge of the plane (thus establishing its two-dimensionality) can be considered “fundamental views,” since they present the most basic structures of the data in the most direct manner.

The MSS Tasseled Cap transformation further orients the data plane such that the two features which define it are directly related to physical scene characteristics. Brightness, the first feature, is a weighted sum of all the bands, and was defined in the direction of principal variation in soil reflectance. It thus measures soil brightness or total reflectance. The second feature, Greenness, is a contrast between the near-infrared bands and the visible bands. The substantial scattering of infrared radiation resulting from the cellular structure of green vegetation, and the absorption of visible radiation by plant pigments (e.g., chlorophyll), combine to produce high Greenness values for targets with high densities of green vegetation, while the flatter reflectance curves of soils are expressed in low Greenness values. A third feature, termed Yellowness, was originally defined in the spectral direction expected to correspond to plant senescence (Kauth and Thomas [2]), but was subsequently redefined to serve as a haze diagnostic (Kauth et al. [3]). As a linear transformation (rotation), the MSS Tasseled Cap transformation preserves the euclidean relationships in the raw data, but captures typically 95 percent or more of the total variability in two readily-interpretable features. Furthermore, since it is an invariant transformation, features are consistent between scenes, and can be so interpreted once the data have been normalized for external effects such as haze level and viewing and illumination geometry. These desirable characteristics have resulted in the widespread acceptance and use of this transformation of MSS data.

In the last several years, anticipation of the now completed launch of Landsat-4, which carries the Thematic Mapper, produced interest in the possibilities of a Tasseled-Cap-like transformation for TM data. The scene reflectance characteristics which result in the planar concentration in MSS data, and the similar placement of some of the TM and MSS bands, suggested that such a transformation might prove useful, while the increase in number and dispersion of spectral bands of the TM as compared to the MSS made data volume reduction and improved interpretability even more attractive.

and present such a transformation based on simulated TM data. That work showed that vegetation and soils TM data in the six reflective bands primarily occupy three dimensions defining two planar data sets, comprising a "transition zone" between the two (Fig. 1). This paper carries that simulation-based transformation to actual TM data, presents a rotation matrix with which to adjust TM data in the six reflective bands to TM Tasseled Cap coordinates, and describes some of the physical processes which cause and are expressed in the TM data structures.

II. MATERIALS AND METHODS

A. Data

The data set used in this analysis consisted of portions of three TM scenes, as described in Table I. All three sub-scenes contain some proportion of agricultural fields, natural vegetation, forest, water, and manmade features (e.g., roads, airports, urban areas). The North Carolina scene, which includes an extensive agricultural holding with unusually large fields surrounded by forest and water, constitutes an ideal scene for derivation of a data transformation, in light of the proximity of a diverse set of cover classes. In addition, because of the long growing season in this locale, a range of crop development stages and conditions was present at the time of data collection, providing, in effect, an entire season of data in a single acquisition. As a result, this scene served as the primary study site in the reported research, with the other two sub-scenes serving to corroborate the results achieved with the North Carolina sub-scene. Samples ranging from 1-in-25 to 1-in-225 were selected, resulting in 9000-13000 pixel samples from the three sub-scenes. These samples constituted the actual spectral data used.

B. Experimental Approach

The process of understanding the structure of TM data, and deriving a TM Tasseled Cap transformation, logically includes both the simulated and actual data analyses, and so the experimental approach encompasses both phases of the analysis. The simulation analysis approach is described by Crist and Cicone [4]. In making the transition from simulated to actual TM data the simulated data transformation was first applied to the actual data, and then adjusted based on the resulting presentations of the data.

The approach taken with regard to understanding the structure of the data, and the expression of physical processes within the data structure, can perhaps best be described as a synthesis of physical understanding (including the spectral characteristics of important scene classes, the expression of particular physical processes in those scene classes, the radiometric characteristics of the sensor, and the interactions of the three) with a degree of intuition based on experience with Landsat MSS data. Corroboration and refinement of the expectations derived by this process were accomplished through analysis of data for particular scene classes in both the simulated and actual data.

III. RESULTS AND DISCUSSION

A. General

Application of the simulation-based coefficients to the actual data resulted in presentations of the data which closely resembled the simulation results, but with slight misalignments of the coordinate axes. A series of small (<<10 degree) rotations applied to the various feature coefficients corrected the misalignments, and resulted in the single transformation matrix which achieved the desired result for all three sub-scenes. This matrix is provided in Table II. Fig. 2 illustrates the three fundamental views of the TM data, transformed to TM Tasseled Cap coordinates, for the North Carolina sub-scene. These results correspond closely to the simulation results shown in Fig. 3. Here as with the MSS data, the "fundamental views" are those which present the data structures in the most direct manner. For TM data from vegetated scenes, these views present: a) the Plane of Vegetation, b) the Plane of Soil, and c) the edges of these two planes.

The three primary features of the TM Tasseled Cap transformation will be discussed first, after which the three fundamental views provided by the transformation will be considered.

B. The TM Tasseled Cap Features

Brightness: The first feature, Brightness, is a weighted sum of all six reflective TM bands (Fig. 4). As such, it is responsive to changes in total reflectance, and to those physical processes which affect total reflectance. Thus differences in soil characteristics such as particle size distribution will be clearly expressed in Brightness, while increases in vegetation density,
which would tend to increase near-infrared response while decreasing visible response, will cause less substantial changes in Brightness.

The naming of the feature reflects its close similarity to MSS Brightness, which is also a weighted sum of all bands. However, because of the influence of the mid-infrared bands, TM Brightness is not identical to MSS Brightness—in the simulation studies, the two had a correlation of 0.77 (Crist and Cicone [5]). Further, MSS Brightness was defined by the primary direction of soil reflectance variation. TM Brightness, on the other hand, is defined by, or defines, the intersection of the two planes as illustrated in Fig. 1. Because soil variation is decidedly two-dimensional in the TM Tasseled Cap space, TM Brightness is not equivalent to the primary direction of soil variability. However, as we will show later, the definition of TM Brightness is comparable to MSS Brightness in terms of Kauth and Thomas's Tasseled Cap "philosophy."

Greenness: The second feature, Greenness is also named to reflect its similarity to MSS Greenness. Both are contrasts between the sum of the visible bands and the near-infrared
CRIST AND CICONE: PHYSICALLY-BASED TRANSFORMATION OF TM DATA—THE TM Tasseled CAP

Fig. 4. TM Tasseled Cap Brightness coefficients.

Fig. 6. TM Tasseled Cap Wetness coefficients.

Fig. 7. Crop spectral development in the Plane of Vegetation view. Point descriptions: (a) bare soil, (b) greening up, (c) full canopy closure, (d) senescence.

Fig. 8. Scene classes in the Plane of Vegetation view.

The coefficients of TM Greenness, illustrated in Fig. 5, are such that the two longer infrared bands essentially cancel each other. Thus only the blue band is left to represent the broader spectral range of the TM, and this band, particularly for green vegetation, is closely correlated to the other visible bands. Thus in the simulation work, TM and MSS Greenness were found to have a correlation of better than 0.99, allowing one to conclude that the two are, for all practical purposes, identical.

Like its MSS counterpart, TM Greenness responds to the combination of high absorption in the visible bands (due to plant pigments and particularly chlorophyll) and high reflectance in the near-infrared (due to internal leaf structure and the resultant scattering of near-infrared radiation) which is characteristic of green vegetation. Greenness (MSS) has been shown to be moderately well correlated with percent canopy closure, leaf area index, and fresh biomass (Bauer et al. [6]). With TM brightness, TM Greenness defines the "Plane of Vegetation," which will be discussed later.

Wetness: Thus far the TM Tasseled Cap features have been comparable to those of the MSS Tasseled Cap—the third TM Tasseled Cap feature represents entirely new information. This feature, which contrasts the sum of the visible and near-infrared bands with the sum of the longer-infrared bands (Fig. 6), has been tentatively named Wetness. While the reason for this choice of names will be more fully explained later, it is in one sense a logical selection, since the longer-infrared TM bands have been suggested to be most sensitive to both soil moisture (Stoner and Baumgardner [7]) and plant moisture (Tucker [8]). One would therefore expect that changes in moisture status would affect these longer-infrared bands more substantially than the visible or shorter-infrared bands, and thus that a contrast of these two sets of bands would highlight moisture-related scene characteristics. Wetness, with Brightness, defines the "Plane of Soils" in TM Tasseled Cap space. In both the simulated and real data analyses, soil moisture status was found to be the primary characteristic expressed in Wetness, although only a limited amount of plant moisture variation was represented in the data sets.

C. TM Tasseled Cap Fundamental Views

The three features just described combine to define two planes, which can be viewed in three fundamental ways. A distinction must be made between a particular view, in which all the data are projected onto two dimensions, and the planes themselves, which only contain specific scene classes. Both will be discussed. Figs. 7-13, used as illustrations in this section, are schematic representations based on actual data.

Plane of Vegetation View: The Plane of Vegetation view, as in Fig. 2(a), contains TM Greenness and Brightness information and is thus, based on the discussion of these features in Section III-B, equivalent to the MSS Tasseled Cap plane. As in the MSS plane, bare soil data points fall in a region along the bottom of the distribution sometimes called the "soil line" (Fig. 7).
As plants begin to emerge in a field, the spectral characteristics of the plants themselves cause an increase in Greenness, while the shadows they cast will, unless the soil itself is very dark, tend to cause a decrease in Brightness (point b in Fig. 7). Continued green vegetation development eventually results in less bare soil or shadowed soil in the field of view, and an increasingly dense vegetation component. During this phase both Greenness and Brightness increase (point e in Fig. 7). The onset of crop senescence results in decreasing Greenness, while the Brightness effect differs with various crops. The crop trajectory in Fig. 7 is, in this respect, must like that of a cereal crop, although the trajectory as depicted is not intended to precisely correspond to any particular crop.

Fig. 8 shows the approximate location of selected other cover classes in the Plane of Vegetation view. All these results are consistent with those obtained in the MSS Tasseled Cap plane, and lend further support to the claim that the TM Plane of Vegetation is in fact equivalent to the MSS Tasseled Cap plane. This equivalency greatly aids the development of our understanding of TM data, since all our insight into the MSS Tasseled Cap plane can be directly applied to this view of the TM data. As a result, we are placed in the position of a) knowing a good deal with regard to scene class responses and the expression of physical scene characteristics in the Plane of Vegetation, and b) being able to more quickly and fully understand the new TM information by using the well-understood Plane of Vegetation as a reference point.

Plane of Soils View: The Plane of Soils view provides one look at the new information present in TM data. In Fig. 2(b) we are looking at the edge of the Plane of Vegetation (lying along the Brightness axis) and at the Plane of Soils “head on.” All the vertical variation in the figure is new information, while the horizontal or Brightness variation is the same as that in the Plane of Vegetation. Fig. 9 shows the same type of crop trajectory as in Fig. 7. In this view, bare soil data may fall anywhere in the data distribution (for example, point a in Fig. 9). As the green plants emerge, shadowing again causes reduction in Brightness, while the vegetation’s spectral contribution drives the data point toward the Plane of Vegetation (the result of increasing near-infrared reflectance changing the balance in the Wetness computation—see Fig. 6). When full canopy closure has been achieved, the crop is found in the Plane of Vegetation (point e in Fig. 9). Additional vegetative development causes variation in this plane toward higher Brightness values since the shaded soil component is largely or entirely removed (point d, Fig. 9). Crop senescence affects Brightness differently for various crops, as mentioned earlier, but the trajectory seems to remain in the Plane of Vegetation until the very end of the development cycle. This is probably a result of an offsetting influence in the reflectance of a senescent canopy—the near-infrared reflectance decreases while the visible reflectance increases, with the net result being fairly constant Wetness values.

Fig. 10 shows the approximate locations of other scene classes in the Plane of Soils view. Here we note that the water and forest classes both appear “wetter” than the other classes. The water result is, of course, intuitively agreeable, while the location of the forested class appears to be the result of shading in the forest canopy. The difference in Wetness values...
between clear and turbid water can be explained by referring to the Wetness coefficients in Fig. 6, and remembering that the reflectance of water in the infrared wavelengths is extremely low. If the infrared bands are taken out of the equation, Wetness becomes a weighted sum of the visible bands, and turbid water, which will by nature of the sediment it carries exhibit higher visible reflectance, will thus appear "wetter" than clear water. Of course, this is an artificial distinction, and the primary point to be made here is that Wetness loses its meaning in distinguishing between water classes or others which exhibit little or no infrared response.

By combining various classes from the simulated and real data sets, a clear direction of moisture variation can be identified. In addition to moist and dry soils, we can treat water as the ultimate wet material, and dry concrete as an extremely dry material. The resultant moisture direction is illustrated in Fig. 11. This direction is obviously correlated to both Brightness and Wetness. Crist and Cicone [4] demonstrate, however, that Brightness alone fails to capture much of the important moisture-related information, particularly with respect to absolute variation in moisture status between soil classes which differ significantly in other characteristics likely to influence Brightness. Such information is contained in the third TM Tasseled Cap feature, and hence the name Wetness is associated with this feature.

Transition Zone View: The final "fundamental view" of the TM Tasseled Cap space is that shown in Fig. 2(c), the Transition Zone view. Here only the edges of the two planes are visible—the Plane of Vegetation along the vertical, and the Plane of Soils along the horizontal. In both the simulated and real data, those samples falling in the Plane of Soils are from bare soil fields, while the samples occupying the Plane of Vegetation are from fully vegetated fields (Crist and Cicone [4]). The data not in either plane are from fields with partial vegetative cover, where both green vegetation and soil are visible to the sensor.

A typical cereal crop trajectory in this view is illustrated in Fig. 12. Starting somewhere on the Plane of Soils (point a in Fig. 12), the addition of green vegetation causes increases in Greenness and Wetness (point b in Fig. 12—see earlier discussion of the two plane views for further explanation). Around the time of complete canopy closure, the trajectory reaches the Plane of Vegetation (point c, Fig. 12). All subsequent green vegetation development, and nearly all of the process of crop senescence, takes place in the Plane of Vegetation—only at the very end of the development cycle does the crop begin to move back out onto the Plane of Soils.

From this description, one can easily understand the reason for the name "Transition Zone." Samples in this spectral region are "in transition" between the Plane of Soils and the Plane of Vegetation. Further, since bare soil and full vegetation are, in this view, confined to perpendicular linear regions, a simple angular measure between the two may serve as an easy means of determining the relative soil/vegetation mix of any particular data point. Various such means of extracting information from this Transition Zone view are currently being investigated.

Fig. 13 shows the same scene classes as previously displayed, this time in the Transition Zone view. Once again we see water and forest at higher Wetness values than the other classes, projected "in front of" the Plane of Vegetation as seen in Fig. 2(a). This projection in the three-space which contains most of the TM data variability suggests that the forest class may be more readily distinguishable from crop or natural vegetation data than it was in the MSS data space.

D. Lesser Features of the TM Tasseled Cap Transformation

The TM Tasseled Cap transformation, as a linear rotation of the six TM reflective bands, includes six features. Fig. 14 illustrates that the vast majority of data variability is captured in the first three features, but some residual variation can be seen in the lesser three features. While the coefficients for these features, as they currently exist, are presented in Table II, most of what will be discussed will be possible shortcomings in their current definition.

"Leakage" from Other Features: Our underlying assumption, based on observation of simulated and real data, is that the TM data occupy three dimensions and define two planes (Fig. 1). The coefficients of the TM Tasseled Cap transformation are intended to rotate the TM data such that Brightness, Greenness, and Wetness exactly correspond to the directions which define the two planes. Graphical analysis of the results of this transformation indicates that this correspondence is very close. If in fact the correspondence is exact, then variation in the lesser features is indeed residual variation above and beyond that seen in the three primary dimensions. It is likely, however, that the correspondence is less than perfect. In this case, some of the variation seen in the lesser features is in fact variation in one or more of the three primary dimensions, and shows up in the lesser features only because of inaccuracies in the applied rotation. Kauth (personal communication) has named this phenomenon "leakage," e.g., some amount of Greenness variation "leaks" into a lesser feature.

As more experience is gained with TM data, adjustment of the TM Tasseled Cap coefficients may be indicated, leading to
increasing precision in the data rotation. In this first version of the transformation, caution must be used in assigning too much significance to variation seen in the lesser features, with the need for caution increasing from the fourth to the sixth feature.

**Physical Association:** A second consideration with these lesser features is their association, or lack thereof, with physical scene characteristics. While the first three features are defined by the structures observed in the TM band space, and are associated with particular physical scene characteristics, the lesser three features are more strongly driven by purely statistical (principal component-like) considerations. Thus while some apparently significant soil-related variation was observed in the fourth feature in the simulation analyses (Crist and Cicone [4]), there may be no clear association between physical processes and the fifth and sixth features. The one remaining degree of freedom represented by these two features may, as a result, be used to define a new feature associated with some other important scene characteristic such as atmospheric haze. This potential is currently being investigated.

**IV. Summary and Conclusions**

In discussing their general approach to understanding MSS data from agricultural regions, Kauth and Thomas [2] wrote:

The first step is to find a variety of ways to display the data. Given the displays, one will notice the structure of the data, will name the data structures, will create some visual model of the data structures. Such a model may suggest a physical interpretation or a quantitative test. One then subjects the description to quantitative tests, resulting in gradual elaboration and refinement of the description, or, in some cases, in its destruction.

In deriving the TM Tasseled Cap transformation, we have first displayed the data, three dimensions at a time, from a variety of perspectives to determine the basic data structure (Crist and Cicone [4]). Through analysis of both simulated and actual TM data (transformed by equivalent but not identical coefficients) from fields about which at least some physical characteristics were known, associations between the transformed features and such physical characteristics were hypothesized, and confirmed through additional data analysis. While the process of gradual elaboration and refinement has only begun, the transformation at this point in time represents a useful tool for interpreting and understanding Thematic Mapper data.

The transformation is not equivalent to a principal components transformation. Principal components analysis can fail to capture the complex structure of TM data (Crist and Cicone [4]), and is extremely scene dependent. The TM Tasseled Cap transformation, on the other hand, specifically emphasizes the inherent data structures, and is intended to be an invariant transformation which can therefore be applied to any TM scene (although atmosphere and illumination geometry will affect results, as may substantial deviation from a mid-latitude, temperate environment). Although to date only three actual TM scenes have been analyzed, the facts that this transformation produced comparable results for all three scenes, separated geographically if not temporally, and that the results were comparable to those achieved with the more diverse simulated data set, suggest that the transformation could indeed be applied directly to other scenes as well. Of course, additional TM scenes should be evaluated to confirm the invariance of the transformation.

As it differs from purely statistical transformations, so the TM Tasseled Cap transformation differs from transformations based purely on a priori assumptions regarding important scene characteristics or data structures. As mentioned earlier, TM Brightness is not the axis of primary soil variation (That axis would roughly correspond to the direction of moisture variation in Fig. 11), but rather is defined by the intersection of the two planes in the TM space. Defining Brightness based on the soil variation alone would result in distortion or loss of vegetation information, since that information is most fully represented in the Plane of Vegetation view. To assume that the directions associated with key physical properties in the spectral space of one sensor will necessarily be duplicated in the spectral space of a different sensor can thus result in overlooking the important and inherent structure of the data. Knowledge of the spectral expression of physical scene characteristics is clearly of central importance, but that knowledge is best used to help understand the structure of the data, rather than to force the data into preconceived structures.

Finally, it should be noted that the "Planes" of Vegetation and Soils are not, in the strict geometric sense, planes, but are rather plane-like three-dimensional data concentrations, even as the MSS "soil line" is not, in the strict geometric sense, a line. In both cases, the geometric terms serve to convey a general idea which facilitates visualization of the actual data structures. Employing the process described by Kauth and Thomas, TM data from the six reflective bands have been found to primarily occupy three dimensions, defining two planes and occupying a zone of transition between the two. The features termed Brightness and Greenness are comparable and equivalent, respectively, to their MSS counterparts, while the third feature, named Wetness, contains new information related at least to soil moisture status. The Plane of Vegetation view provides an equivalent data plane to the MSS Tasseled Cap plane, while the Transition Zone view may provide improved ability to determine the relative mix of vegetation and soil components in the sensor field of view.

The "fundamental views" of the data provided by the TM Tasseled Cap transformation provide a means by which the spectral expression of physical processes, and various ratios or other derived features, can be better understood. It represents no alteration of the basic data, but simply a change in viewing perspective, a rotation which presents the TM data in a more accessible fashion. By concentrating the vast majority of data variability in three features, the transformation serves a data volume reduction function which should be of particular value in the processing of large regions and/or multiple acquisitions. Finally, the direct link between features and physical scene characteristics enhances both the interpretation of observed spectral variation and the prediction of the spectral effects of particular changes in scene characteristics.

**Acknowledgment**

The authors wish to express their appreciation to R. J. Kauth of ERIM, who contributed both through his pioneering work on the MSS Tasseled Cap transformation and through dis-