Reprinted from

Eighth International Symposium

Machine Processing of

Remotely Sensed Data

with special emphasis on

Crop Inventory and Monitoring

July 7-9, 1982

Proceedings

Purdue University
The Laboratory for Applications of Remote Sensing
West Lafayette, Indiana 47907 USA

PROFILE MODELING FOR CROP DISCRIMINATION

G.D. BADHWAR

National Aeronautics and Space
Administration/Johnson Space Center
Houston, Texas

ABSTRACT

This paper presents a methodology to extract physical and morphological characteristics of crop type from the multitemporal-multispectral Landsat multispectral scanner data. It also describes how these characteristics can also be calculated using whether information and crop calendar model and thus provide a truely automated method of identifying crop types. Moreover, because of the established link to crop morphology, these features are expected to hold in foreign countries also.

I. PROFILE MODELING FOR CROP DISCRIMINATION

It is a basic characteristic of living things that they undergo birth and death, growth and decay, change and transformation. They are, therefore, involved in dynamic process of development in time. This aspect of change with time usually leads immediately to the formulation of differential equations. In nature most living things (plants) do have a propensity for geometrical growth which is held in check by a sufficient degree of competition for food, death and destruction. A typical plant population grows rapidly at an increasing rate if it starts in an environment with adequate food supply. As time elapses, the food supply becomes less adequate and the death rate increases. This curve of population growth that is typical of biological systems, first suggested by Verhulst, first goes as an S-shaped and then declines as an S-shaped curve. The combined curve is bell shaped.

A number of studies have suggested a strong correlation between the green vegetation and a linear transformation of Landsat MSS data, the Kauth-Thomas Greenness.[1] This greenness function, $\rho(t)$, is continuous as a function of time, t, and has a bell shaped curve called a profile. A heuristic model of this time behavior was suggested to be of the form[2,3]

$$\rho(t) \alpha t^{\alpha} e^{-\beta t^2} \qquad (1)$$

where α and β are crop and condition related constants. This function is highly peaked at time $t_p = \sqrt{\alpha/2\beta}$ and provided adequate representation of the Landsat spectral data. With the availability of rather frequent temporal-spectral measurement on ground level plot data, Bauer et al.[4] observed that the corn greenness profile has a flat top in time, a characteristic that cannot be produced by model of equation (1). Crist[5], again on completely empirical grounds, suggested a new model that rectifies the deficiency of model (1). However, it does not provide any understanding into the nature of this profile behavior. We suggest that original idea of Verhulst can be modified and can provide an explanation of these profiles.

The Kauth greenness is approximately the difference in the reflected amplitude of light in the 0.7-0.8 μm and the part containing the chlorophyll band 0.6-0.7 μm. In a living annual crop plants new leaves are being formed and old leaves die off continuously with time up to certain time when no new leaves are formed and old leaves continue to die off. It is also well known that because of the very high absorption coefficient in the chlorophyll band of 0.6-0.7 μm only the upper few layers (about two) of leaves are responsible for the returned signal whereas six to eight layers of leaves from canopy are responsible for the returned light amplitude in 0.7-0.8 μm region. Since the leaves die-off rate depends on their location up from the soil, the reflected amplitude in 0.7-0.8 μm responds to the average die-off rate. We thus postulate that the greenness responds to the net rate of the birth rate of leaves--death

rate of leaves. We also treat the greenness as a continuous variable, which is accurate enough if the plant population density is large. (This is a minor restriction.) In addition to the birth rate and death rate, there are other retardation factors that would not permit the greenness to rise indefinitely, for example, the availability of food.

Let us suppose that the average rate of growth of greenness under favorable conditions is k per individual plant (that is k(t) = birth rate - death rate), so that in time dt there is an increase of $k(t)\rho dt$ in the greenness. This means that $d\rho = k\rho dt$. In addition to birth-death process, there are other retardation factors, such as availability of food, i.e., that the net growth rate is not $\rho(t)$ but $\rho - r\zeta(\rho)$ where r is the retardation constant. The basic differential equation is now

$$\frac{d\rho(t)}{dt} = k\rho(t) - r\zeta(\rho) \qquad (2)$$

The simplest assumption to make is that the function $\zeta(\rho)$ is ρ^2; i.e., the retardation rate per individual plant is proportional to the population size

$$\frac{d\rho(t)}{dt} = k\rho(t) - r\rho^2 \qquad (2a)$$

with the solution

$$\rho(t) = \frac{\rho_m}{1 + (\frac{\rho_m}{\rho_o} - 1)\, e^{-\int_{t_o}^{t} k(t)dt}} \qquad (3)$$

where at $t=t_o$, $\rho(t)=\rho_o$ and ρ_m is the maximum ("saturation") value of greenness.

Figure 1 is fit of the model form (3) to the soybean field measurement data of Bauer et al.[4] assuming k(t) = constant. The fit to the data is good and the calculated value of t_o is completely consistent with the observed emergence date. Rood and Major[6] have measured the number of leaf and the leaf development rate of early maturing maize (Zea may L.). From their measurements one can deduce that the absolute value of $k(\equiv k_2)$ after peak greenness is approximately twice the value of k before the peak greenness (Figure 1). A simple functional form[7] that satisfies this constraint is $k(t) = \{1-\exp(-a(t-t_p))/1 + 2\exp(-a(t-t_p))\}$ where t_p is the time of peak and a is constant that controls the transition zone when k_1 to k_2. A simple integration of this function and substitution in equation (3) gives the form of $\rho(t)$ which is bell shaped curve. It is by no means implied that k(t) of the form suggested above is indeed the correct form; it does suggest, however, that different forms of k(t) can produce a bell shaped or a flat top curve. Experiments, such as that of Rood and Major[6] and good detailed and frequent measurement of spectral data are necessary to establish the correct form of k(t). Until the precise form of k(t) can be determined from field experiments, approximations to model (3) that adequately describe the Landsat greenness behavior would have to suffice. The form suggested by Crist[5] has five free parameters excluding the bare soil greenness and is thus impractical from an operational point of view. The invariant form of model (2) given by

$$\rho(t) = \rho_o + (\rho_m - \rho_o)\left(\frac{2\beta e}{\alpha}\right)^{\alpha/2} (t-t_o)^{\alpha} e^{-\beta(t-t_o)^2} \qquad (4)$$

has only three free parameters, again excluding the soil greenness ρ_o. In either this model, or the model proposed by Crist[5] the assumption of a constant ρ_o is well supported by independent experimental data. This assumption will be made throughout the remainder of this paper. The parameter, t_o, is emergence date of the crop. The function (4) has two inflection points t_1 and t_2 and peaks at time t_p. These are given by

$$t_p - t_o = \sqrt{\alpha/2\beta}$$

$$t_1 - t_o = \frac{(2\alpha+1) - \sqrt{8\alpha+1}}{4\beta}^{1/2}$$

$$t_2 - t_o = \frac{(2\alpha+1) + \sqrt{8\alpha+1}}{4\beta}^{1/2}$$

and the time distant between the two points t_2 and t_1 is given by

$$\sigma^2 \equiv (t_2-t_1)^2 = \frac{1}{2\beta} + \frac{\alpha}{2\beta}\left[1 - \left(1 - \frac{1}{\alpha}\right)^{1/2}\right] \simeq \frac{1}{\beta}$$

The parameter $\sqrt{\beta}$ is thus effectively the width of the greenness profile. This profile is thus characterized by its maximum value, ρ_m, at time t_p-t_o, the width, σ, and a slope parameter α. This model thus converts multidimensional spectral and temporal data to three variables and can lead to simple classifier design. Investigations have shown that σ is effectively the grain filling period, which is a crop characteristic. These parameters of the model are crop growth varaibles and can be expected to be stable over large geographic areas and can be predicted from agromet models of crop growth.

Figure 2 shows a scatter plot of a subset of ground truth pure pixels selected from a systematic grid of 418 dots. The two axis are the calculated

value of ρ_m and σ obtained by fitting each pixel with the model of equation (4). The symbol S is for pixels called soybean using ground truth, the symbol C is for pixels called corn, and the symbol * is for all pixels which are neither corn or soybeans. Values of σ>100 days are plotted at day 100. It is quite apparent that a simple cut at, say, σ>90 days would reject all other crop categories and keep only pure corn and soybean pixels. Having rejected these pixels, it is also clear that there is very good separation between corn and soybeans and a simple linear discriminant would be sufficient to separate these two crops. With some rather straightforward modifications to this idea an end-to-end near-harvest classification approach has been formulated and evaluated over a large geographic area in 3 years (1978, 1979, and 1980).

The classification approach basically consists of converting the multitemporal Landsat measurements into Kauth greenness and fitting it to mdoel form (4) and calculating ρ_m, σ, α, t_p for each pixel. Using a simple gradient model pure pixels are selected from a systematic grid of 836 pixels spread out over the entire segment. These pixels are labeled automatically using the basic idea presented in the previous paragraph. The labeled pixels are then used to train the Ho-Kashyap[8,9] linear classifier. This classifier determines the separation plane in the four-dimensional feature space. Using these separation planes each pixel is classified into one of the three categories of corn, soybean, or other. In order to set more precisely the cuts on σ and restrictions of the ratio of ρ_m of soybean to ρ_m of corn, four U.S. Corn Belt segments (two each in 1978 and 1979) were used.

A. DATA ANALYSIS

The data set consisted of a total of 56 5 nm-by-6 nm segments spread over 3 years. There were 40 1978 segments with a geographic distribution from Minnesota in the north to Mississippi in the south and extending from eastern Indiana to western Nebraska. The selection of these segments is discussed in greater detail elsewhere[10]. In addition, nine 1979 segments in the U.S. Corn Belt and seven 1980 segments in Iowa were selected for processing. The available acquisitions on these segments were visually screened so as to have minimal cloud cover and no multitemporal misregistration problems. Only acquisitions that fall within the time span of calendar day 120 to 273 were used for analysis. A basic restriction of this model form is that there must be at least one acquisition on either side of the time of peak greenness, which occurs around the blister stage for corn. Clearly, for good results one requires acquisitions that are spread over the entire growing cycle of the crop; however, only three well-distributed acquisitions are all that is necessary.

B. RESULTS

The data set of these 56 segments was divided into various groups to look at geographic and year-to-year effects. A subset of 27 1978 and 1979 segments in the U.S. Corn Belt were first selected. These segments are the same set that were processed by procedure CS1-B[11]. Figure 3, a scatter plot of the difference between the classification determined proportion and the true ground truth determined proportion versus the ground truth proportion for corn, soybean, and other. One finds no particular dependence of the proportion error as a function of ground truth proportion. Table 1 gives a summary of the relevant statistics and indicates a relative mean error of <2% for corn and soybean. A comparison of these results with those of Meltzer[11] show that on the average these results are about a factor of three better.

A data set of 40 1978 segments was selected from the set of 56 segments. These segments contain many segments that fall outside the main U.S. Corn Belt region. Figure 4 and Table 2 provide the relevant statistics on these segments. Both the bias and standard deviation of the bias are somewhat higher than the results shown in Table 1. However, these results are quite comparable to the results reported on these segments in an earlier study using ground truth derived levels[10]. This indicates that the feature derived from the profile ρ_m and σ can be extended over a large geographic area. A number of previous attempts to extend crop signatures had failed because one was not extending crop growth variables but only spectral appearances.

Figures 4 and 5 give the results of an analysis of nine 1979 segments and seven 1980 segments. Tables 3 and 4 summarizes the statistics on these segments. A comparison of the 1978, 1979, and 1980 statistics shows no year effect. Figure 6 and Table 5 gives the results of all 56 segments. The results are based on largest test so far and show little or no bias.

C. CONCLUSIONS

A complete automatic unsupervised classification approach to estimate the corn, soybean, and other proportion in a Landsat 5 nm-by-6 nm segment has been developed. The technique involves extracting crop growth variables that are predictable from agromet models and provide a signature extension not possible before. The technique has been applied in 3 years (1978, 1979, and 1980) and shows no year effect as well as any geographic effect. The results to date are by far the best of any other technique currently available. The technique can be improved to reduce the bias however.

II. REFERENCES

1. Kauth, R.J. and G.S. Thomas, 1976. The tasselled cap, A graphic description of the spectral-temporal development of agricultural crops as seen by Landsat Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette, IN, 4B, 42.

2. Badhwar, G.D., 1980. Crop emergence data determination from spectral data. Photogram. Eng. Remote Sens. 46:369-377.

3. Crist, E.P. and W.A. Malila, 1980. Temporal-spectral analysis technique for vegetation applications of Landsat. 14th International Symposium on Remote Sensing of the Environment, San Jose, Costa Rica, April 23-30, 1980, 2:1031-1040.

4. Bauer, M.E., L.L. Biehl, C.S.T. Daughtry, B.F. Robinson, and E.R. Stoner, 1979. Agricultural scene understanding and supporting field research, p.24-26. Final Rep., Contract NAS9-15466. Lab for Appl. of Remote Sensing, Purdue University, West Lafayette, Ind.

5. Crist, E.P., 1982. Cultural and environmental effects on the spectral development patterns of corn and soybeans, Field data analysis. ERIM Technical Report SR-E2-04224.

6. Rood, S.B. and D.J. Major, 1981. Diallel analysis of leaf number, leaf development rate, and plant height. Crop Science 21, 867-873.

7. Goel, N.S. and G.D. Badhwar, 1982. Unpublished work.

8. Ho, W.C. and R.L. Kashyap, 1965. Algorithms for linear inequalities and its applications. IEEE Trans. Elec. Comp., EC14:683-688.

9. Ho, W.C. and R.L. Kashyap, 1966. A class of iterative procedures for linear inequalities. J. Saim Control 4:112-115.

10. Badhwar, G.D., J.G. Carnes, and W.W. Austin, 1982. Use of Landsat-derived temporal profiles for corn-soybean feature extraction and classification. Remote Sensing of Environment 12, 57-79.

11. Meltzer, M., 1982. Evaluation of C/SI-B Corn-soybean procedure. Proceeding of Quarterly Interchange Meeting (ITD)--Johnson Space Center, Houston, Texas, March 1982.

Table 1. Results of 1978 and 1979 Segments.

	CORN	SOYBEAN	OTHER
$\bar{e}$ (ERROR)	-0.77	-0.48	1.28
$S_{\bar{e}}$ (S.D.)	6.55	4.79	6.04
RME	-1.88	-1.74	4.05
MAE	5.74	3.79	4.96
AVE. PROP. %	40.98	27.43	31.60
NO. OF SEG.	27	27	27

Table 2. Results of 1978 Segments

	CORN	SOYBEAN	OTHER
$\bar{e}$ (ERROR	1.41	-0.50	-1.28
$S\bar{e}$ (S.D.)	8.21	6.30	7.96
RME	3.95	-1.97	-3.28
MAE	6.31	5.00	5.49
AVE. PROP %	35.75	25.33	38.92
NO. OF SEG.	40	40	40

Table 3. Results of 1979 Segments

	CORN	SOYBEAN	OTHER
$\bar{e}$ (ERROR)	-1.42	-2.25	3.66
$S\bar{e}$ (S.D.)	8.47	4.71	9.21
RME	-3.29	-7.06	14.60
MAE	7.75	3.75	6.30
AVE. PROP. %	43.07	31.88	25.09
NO. OF SEG.	9	9	9

Table 4. Results of 1980 Segments

	CORN	SOYBEAN	OTHER
$\bar{e}$ (ERROR)	2.66	1.43	-4.15
$S_{\bar{e}}$ (S.D)	8.65	2.21	8.14
RME	7.60	7.92	-9.30
MAE	6.77	1.91	7.02
AVE. PROP %	34.96	18.11	44.61
NO. OF SEG.	7	7	7

Table 5. Results of All Segments

	CORN	SOYBEAN	OTHER
$\bar{e}$ (ERROR)	1.11	-0.54	-0.84
$S_{\bar{e}}$ (S.D)	8.24	5.74	8.03
RME	3.02	-2.11	-2.25
MAE	6.60	4.11	5.81
AVE. PROP %	36.83	25.48	37.41
NO. OF SEG.	56	56	56

Figure 1. A plot of observed Kauth-Thomas greenness reflectance of a soybean plot observed by Bauer et al., as a function of the observation date. The solid curve is the fit to the sigmoidal curve of model form (3) with k(t)=k, a constant.

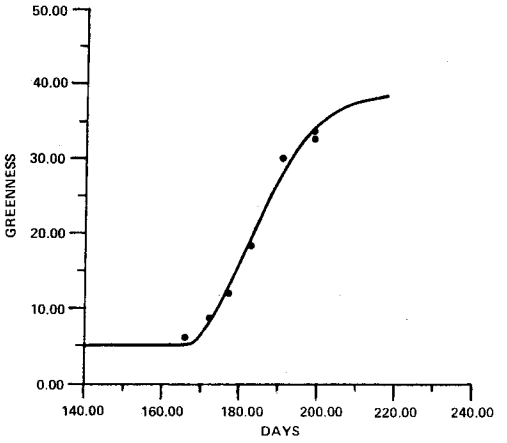


Figure 2. A scatter plot of the maximum value of greenness and the width in days between the two inflection points of a selected subset of pixels in segment 0886 in 1979. C is a corn pixel, S is a soybean pixel and * is pixel that is other corn or soybean. Value of $\sigma = t_1 - t_2$ greater than 100 days are plotted at 100 days.

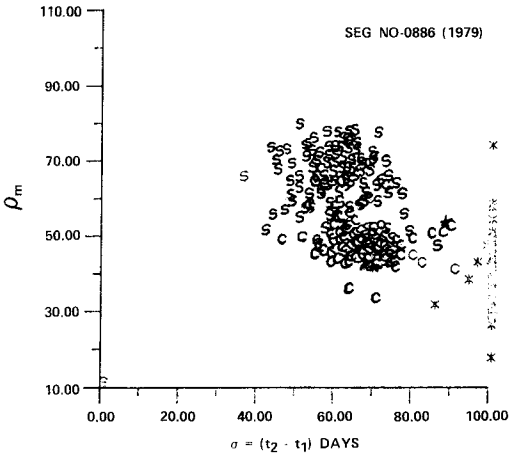


Figure 3. A plot of the error in proportion estimate from classifier, defined as the difference between proportion estimate from classifier and ground truth proportion, and 'true' ground truth proportion for 27 segments used in study by Meltzer[11].

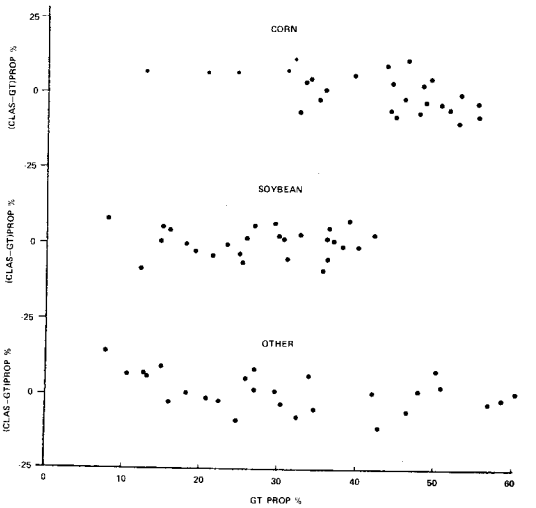


Figure 4. Same plot as in figure 3 except it is for 40 segments in 1978 used in earlier study by Badhwar et al.[10].

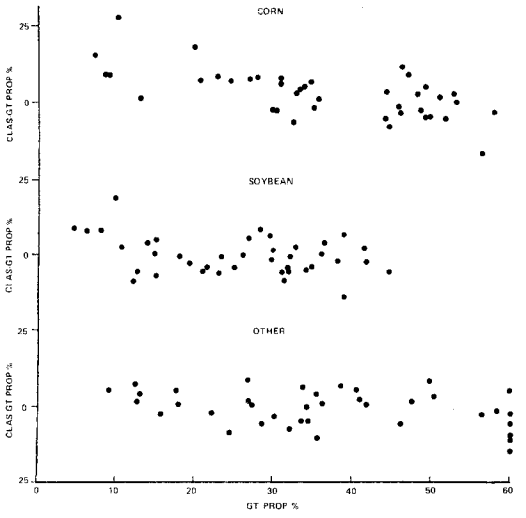


Figure 5. Same plot as in figure 4 but for nine segments in 1979.

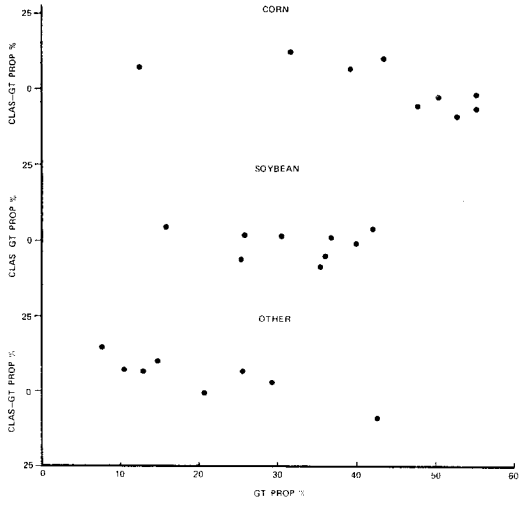


Figure 6. Same plot as in figure 4 but for seven segments in 1980.

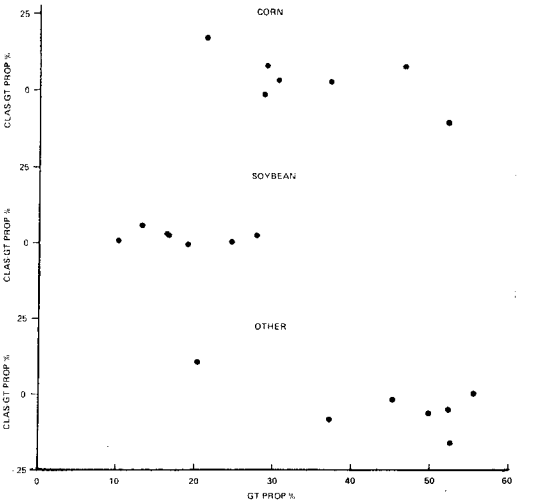
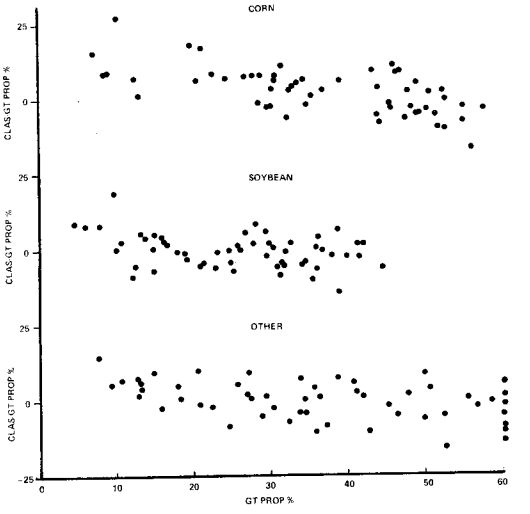


Figure 7. Same plot as in figure 4 but for all 56 segments given in figures 4, 5, and 6.



Gautam D. Badhwar. Received B.S.-1959 in Physics from Agra University; received Ph.D.-1967 in Physics from University of Rochester. Dr. Badhwar has worked as a Research Assistant at the Tata Institute of Fundamental Research in Bombay, India; as an Assistant Professor of Physics at the University of Rochester; as a Senior National Academy of Science Fellow and as a Physicist in the Space Physics Division of NASA/JSC in experimental and theoretical research in cosmic radiation; and as a Physicist in the Earth Resources Research Division of NASA/JSC doing research in agricultural remote sensing.