Information Note 123068
Purdue University

Some Notes on Use of PICTOUT Program

by

Roger M. Hoffer

PICTOUT is an extremely useful program for locating various fields or areas of interest in the multispectral data. A number of questions on the most efficient manner of using this program have arisen. These questions are strictly from the agricultural classification analysis standpoint, and do not involve use of PICTOUT for data quality analysis.

One of the questions seem to be the size of the area and the number of samples required to generate the best gray-scale printout.

In an attempt to answer this question, a small test area (A), approximately 2 pages in length containing 123 lines and 110 sample lines was histogramed. Every line/every sample was used in this area. A second area (B), 587 lines in length, in which every other line/ every other sample was used resulted in 294 actual lines of data and a total of about 32,000 sample points. The third area was the same as (B) -- 587 lines, but in this case every line/every sample was used, giving a total of about 129,000 sample points, or four times the number in Group (B).

Table 1 gives us the times required to calculate the histograms for the various areas and when 2, 4, 6 and 12 wavelength bands were used to calculate the histogram for these areas. The pertinent points to be noted here is that <u>increasing the number of wavelength bands does increase the time required to do the histogram but not by a</u>

great amount. However, increasing the number of sample points considerably increases the time required to do the histograms as shown for Data Set B (every other line/every other sample point) as compared to Data Set C (every line/every sample), both sets of data covering the same ground area.

In comparing the results of the display it was found that Data Sets B and C produced almost identical histograms in both channels 9 and 12. In Channel 12, the display was exactly the same for both sets of data. In Channel 9, starting with symbol M, there was a one point difference in the histogram displays, which resulted in a nearly identical display for the two sets of data. Conclusion: For obtaining histograms and pictorial displays of relatively large areas, every other line and every other sample (rather than every line and every sample) should be used to obtain the histogram. The results will be very similar and a relatively large amount of time is saved (although the time is still only a matter of a few minutes).

The comparison between the large area, Data Set B (or C), and the small area (Data Set A) showed relatively minor differences. In many cases one display showed a better tonal difference than the other display due to the differences in symbols being used, rather than in the way the gray levels were grouped. In some cases, however, the differences in the way the gray levels were grouped did result in quite distinct differences in the displays. In some cases one display seems to be better and in other cases the other display seems to be better.

Conclusion: If detailed information is required for a small area and the tonal differences are rather subtle, it would appear that one should obtain a histogram for that small area alone, and

then get the gray scale printout based upon the histogram for the small area. For general use in finding various locations in the data, the histogram should be from a relatively large area containing all of the materials of interest in the flight lines.

Additional Conclusions and Recommendations

In the first attempt at locating various areas of interest in the data, it is recommended that the histogram be based on a relatively large area. This large area hopefully will include materials of all different multispectral signatures which will be encountered in the data. To save time in calculating the histogram it is recommended that every other line/every other sample be used, rather than every line/every sample. Unless you are sure that you do not need more than two wavelength bands for display purposes, it is recommended that the histograms be calculated in 4-7 wavelength bands, and then a separate card be prepared to display only the two wavelength bands of data believed to be most important. If other wavelength bands are needed, the displays are easily obtained without having to recalculate the histograms.

If a great deal of difficulty is encountered in locating particular areas of interest because of subtle differences in tonal response or because of the differences in the symbols being used to display the data, a histogram and display can be obtained for the small area of interest. After that, a person may want to display the data using specially designated symbols or gray level groupings, rather than those set or calculated by the computer. However, for most purposes these procedures are not recommended.

The user seldom needs to make use of the capability of PICTOUT to graph columns or graph lines. It is well to know that this

capability exists, but unless a particular study is being done comparing scanner data with microdensitometer studies, or unless a need exists to determine the relative response at a particular sample point, I see little need for the user to make use of the graph lines or graph columns capability. The user generally does not need to use the histogram display unless he intends to set his own levels or symbols for the gray-scale printouts, in which case it is necessary to obtain a histogram of data in the channels of interest.

The histogram display capability should be used primarily in conjunction with HD (XXX), where you plot 100 bins of data starting at a designated bin, and with each selected channel on a separate axis. Unless you check it out first, it is recommended that you obtain the display starting with bin 120. A simple way of checking this would be to obtain a histogram of requested channels on the same axis using test character HD. This shows all of the channels of interest on the same graph and this will show the maximum and minimum levels of response for each channel of interest.

In this study, Channel 12 (.8-1.0μ) was <u>much</u> better than Channel 11 (.72-.80μ) in showing distinct tonal differences between materials such as bare soil, green vegetation and water. This leads to the recommendation that PICTOUT be changed to allow Channels 9 and 12 to be used as default options. I would also recommend that every other line and every other column be used in the default option for the histogram calculation control cards.

Gray-scale printouts illustrating the results discussed in this information note will be filed in the analysis results file.

Table 1

A - Data Used

Data Set	Total Lines Involved	Every Line, Every Sample	or	Every Other Line, Every Other Sample	Total No. Sample Points Used
A	123	X			13,530
B	587			X	32,340
C	587	X			129,360

B - Times Required to Calculate Histograms

No. of Wavelength Bands \ Data Set	Time		
	A (13,530 sample points)	B (32,340 sample points)	C (129,360 sample points)
2	15 sec	55 sec	2 min 0 sec
4	21 sec	1 min 5 sec	2 min 38 sec
6	25 sec	1 min 14 sec	3 min 14 sec
12	35 sec	1 min 45 sec	5 min 3 sec