COMPUTER GRAPHICS AND IMAGE PROCESSING 4, 172–179 (1975)

SHORT NOTE

Texture Analysis Using Gray Level Run Lengths

MARY M. GALLOWAY

Computer Science Center, University of Maryland,
College Park, Maryland 20742

Communicated by A. Rosenfeld

Received October 15, 1974

A set of texture features based on gray level run lengths is described. Good classification results are obtained with these features on a set of samples representing nine terrain types.

A wide variety of features have been used for visual texture analysis; see [*1*] for a review. Some of these feature sets have included features based on gray level run lengths, but run length features have not been used extensively. This note reports on the use of such features to classify a set of terrain samples that have also been studied by Haralick [*2*].

A gray level run is a set of consecutive, collinear picture points having the same gray level value. The length of the run is the number of picture points in the run. For a given picture, we can compute a gray level run length matrix for runs having any given direction. The matrix element (i,j) specifies the number of times that the picture contains a run of length j, in the given direction, consisting of points having gray level i (or lying in gray level range i). The following example shows a 4×4 picture having four gray levels (0–3) and the resulting gray level run length matrices for the four principal directions.

Picture:

0	1	2	3
0	2	3	3
2	1	1	1
3	0	3	0

Run length:

Gray level	0°	1	2	3	4
	0	4	0	0	0
	1	1	0	1	0
	2	3	0	0	0
	3	3	1	0	0

45°	1	2	3	4
0	4	0	0	0
1	4	0	0	0
2	0	0	1	0
3	3	1	0	0

Gray level	90°	1	2	3	4	135°	1	2	3	4
	0	2	1	0	0	0	4	0	0	0
	1	4	0	0	0	1	4	0	0	0
	2	3	0	0	0	2	3	0	0	0
	3	3	1	0	0	3	5	0	0	0

Computation of these matrices is very simple. The number of calculations is directly proportional to the number of points in the picture. Also, the entire picture need not reside in core. Only two rows of picture values are needed at any one time to compute the matrices.

To obtain numerical texture measures from the matrices, we can compute functions analogous to those used by Haralick [2] for gray level co-occurrence matrices. Let

$p(i,j)$ be the (i,j)th entry in the given run length matrix;

N_g be the number of gray levels in the picture,

N_r be the number of different run lengths that occur (so that the matrix is N_g by N_r), and

P be the number of points in the picture.

1. Short Runs Emphasis

$$\mathrm{RF1} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{p(i,j)}{j^2} \Big/ \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j).$$

This function divides each run length value by the length of the run squared. This tends to emphasize short runs. The denominator is the total number of runs in the picture and serves as a normalizing factor.

2. Long Runs Emphasis

$$\mathrm{RF2} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} j^2 p(i,j) \Big/ \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j).$$

This function multiplies each run length value by the length of the run squared. This should emphasize long runs. The denominator is a normalizing factor, as above.

3. Gray Level Nonuniformity

$$\mathrm{RF3} = \sum_{i=1}^{N_g} \left(\sum_{j=1}^{N_r} p(i,j) \right)^2 \Big/ \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j).$$

This function squares the number of run lengths for each gray level. The sum of the squares is then divided by the normalizing factor of the total number of runs in the picture. This should measure the gray level nonuniformity of the picture. When runs are equally distributed throughout the gray levels, the function

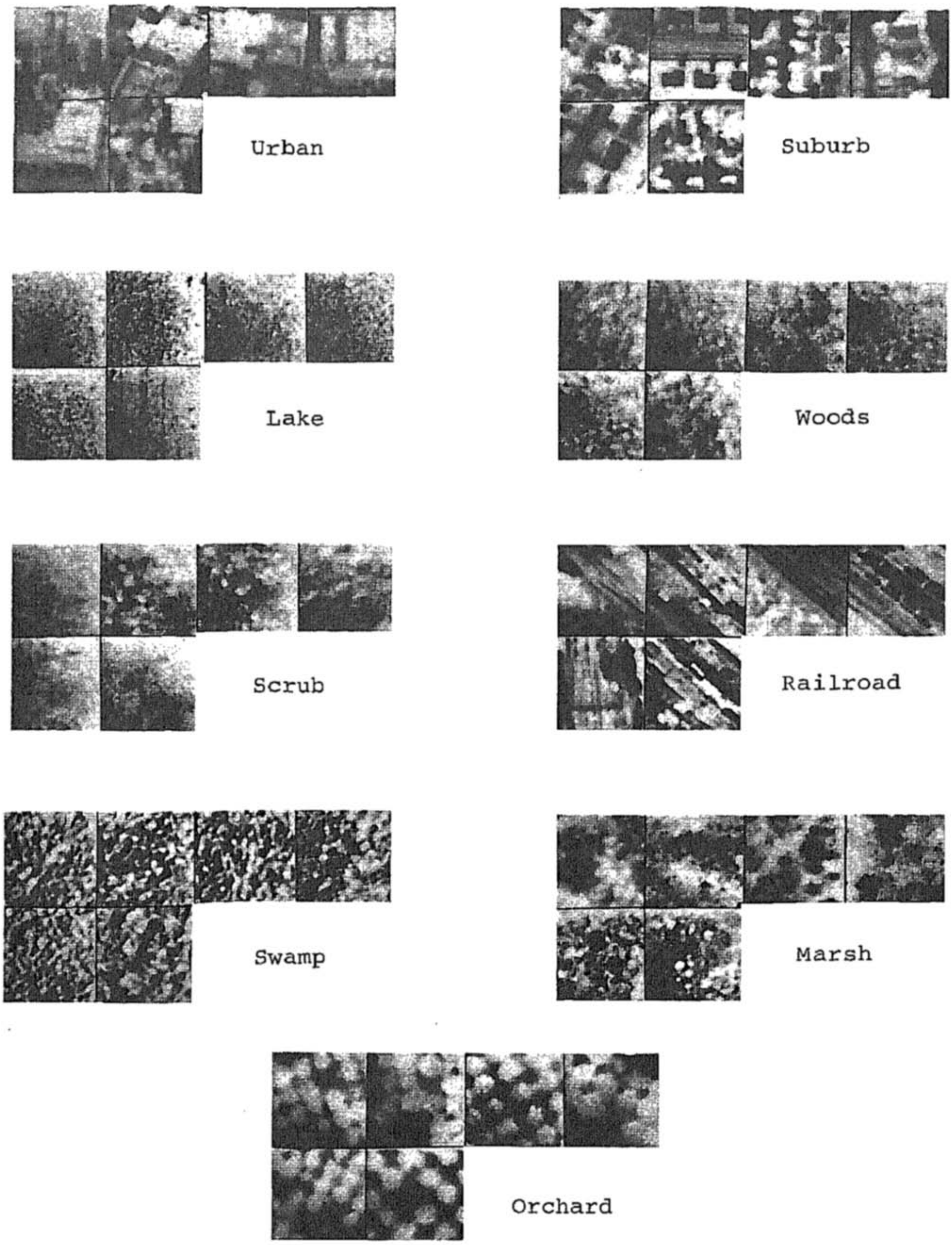

FIG. 1. Terrain samples used in texture analysis study.

takes on its lowest values. High run length values contribute most to the function.

4. Run Length Nonuniformity

$$\mathrm{RF4} = \sum_{j=1}^{N_r} \left(\sum_{i=1}^{N_g} p(i,j) \right)^2 \Big/ \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j).$$

This function squares the number of runs for each length. The sum of the squares is then divided by the normalizing factor. This function measures the nonuniformity of the run lengths. If the runs are equally distributed throughout the lengths, the function will have a low value. Large run counts contribute most to the function.

5. Run Percentage

$$\mathrm{RF5} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i,j)/P.$$

This function is a ratio of the total number of runs to the total number of possible runs if all runs had a length of one. It should have its lowest value for pictures with the most linear structure.

The run length features just described were measured for the set of 54 terrain samples shown in Fig. 1. The samples are taken from nine categories: orchard, wood, urban, suburb, lake, marsh, swamp, railroad, and scrub.

Each picture was digitized into a 64 by 64 array with gray levels ranging from 0 to 63. Gray scale normalization was performed to make all the pictures have a uniform gray level frequency distribution [*2,3*]. Before the run length matrices were computed, the gray levels were grouped into eight sets. That is, the runs were computed for eight gray level groups: 0–7, 8–15, 16–23, 24–31, 32–39, 40–47, 48–55, and 56–63. The run lengths were also grouped into the ranges 1, 2–3, 4–7, 8–15, 16–31, and 32–64. Therefore, the run length matrices were 8×6 arrays containing 8 gray level groups and 6 run length groups. The run length function values for these matrices can be found in [*4*].

To illustrate the classification power of the functions, five typical graphs are shown in Figs. 2–6. These graphs plot one function value against another function value for all the pictures. In most instances, the best results were obtained by plotting different directions of the same function, or the same direction of two different functions against each other. The five graphs shown here were randomly chosen as a representative sample of the 190 possible graphs. They are defined as shown in Table 1. The symbols shown on the graphs correspond to the nine terrain types: B = suburb, C = scrub, L = lake, M = marsh, O = orchard, R = railroad, S = swamp, U = urban, W = woods. An equals sign means that two or more samples should have been plotted at the same point.

From an inspection of the graphs, six of the categories separate from the others for at least five out of the six samples in that category (83% accuracy). These six categories and the graphs which separate them are shown in Table 2. All of the categories can be separated if the graphs are analyzed in a step-by-step

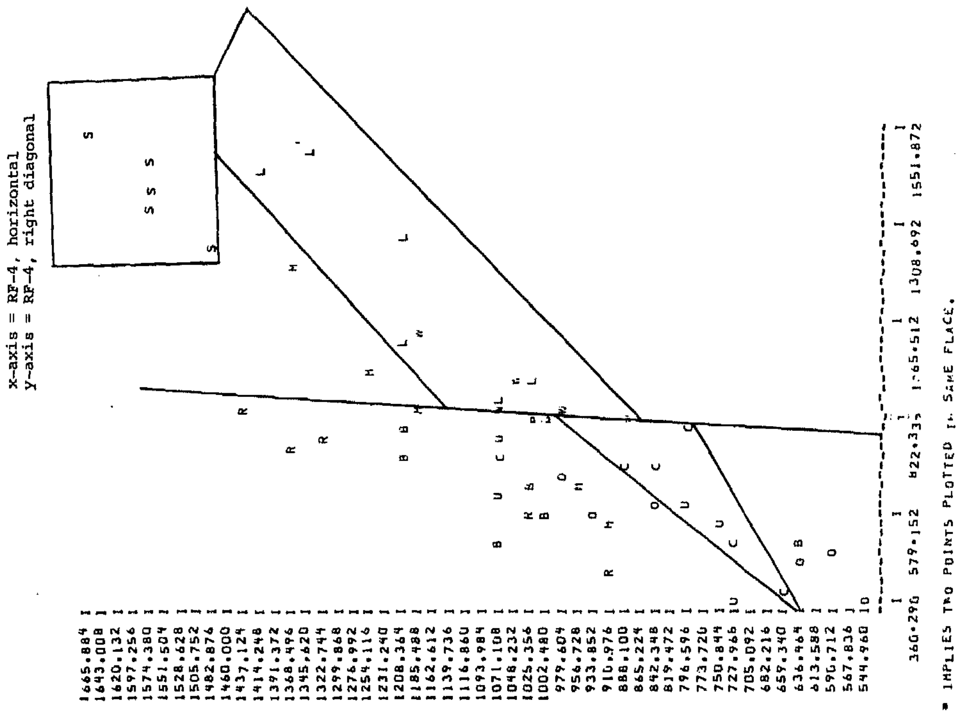

FIGURE 3

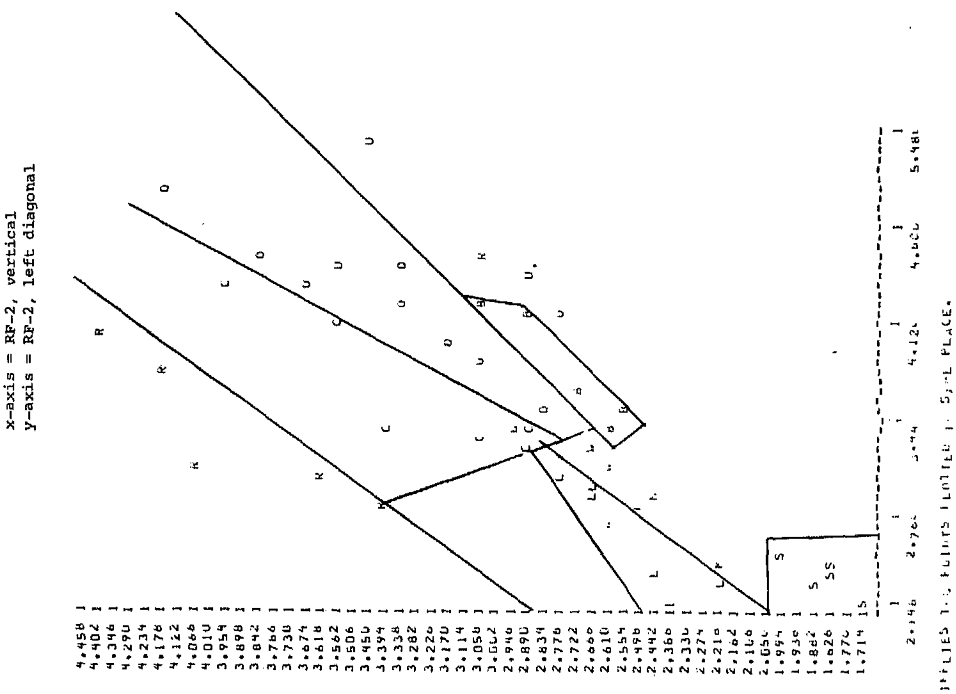

FIGURE 2

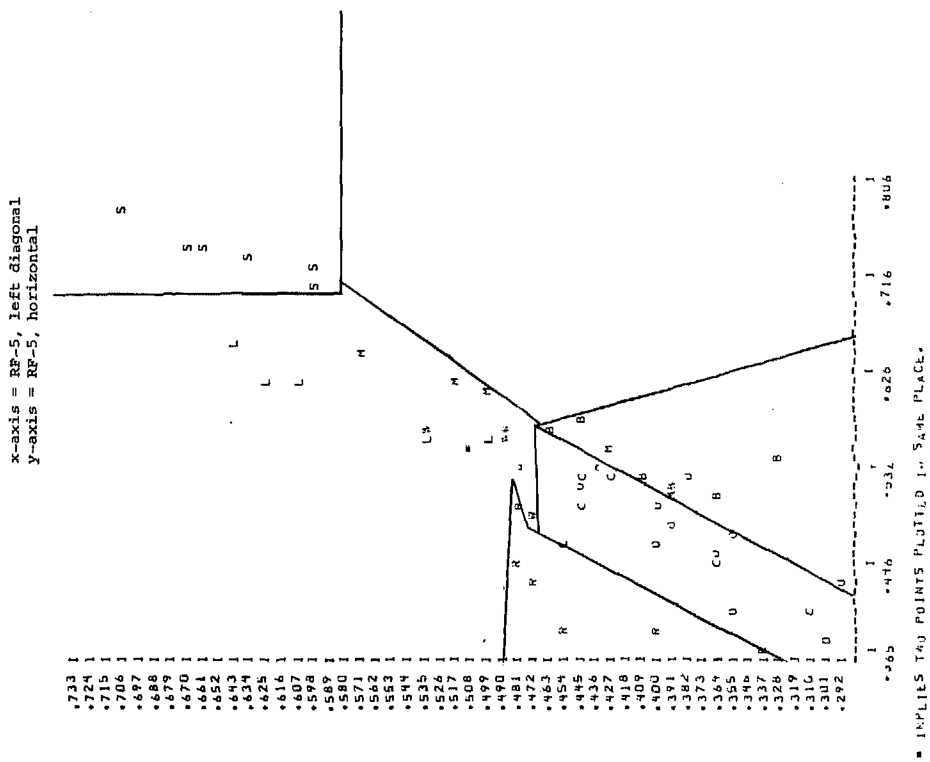

FIGURE 4

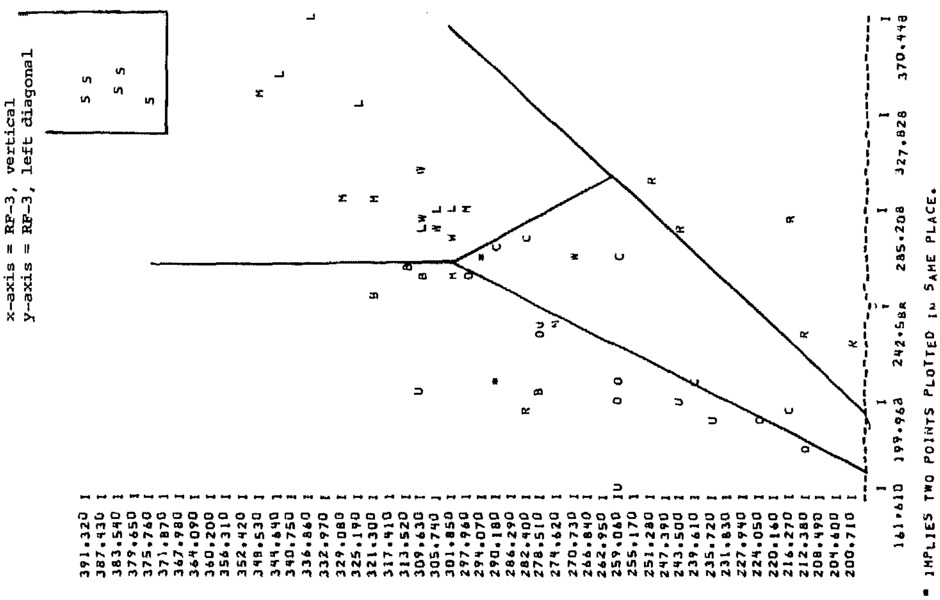

FIGURE 5

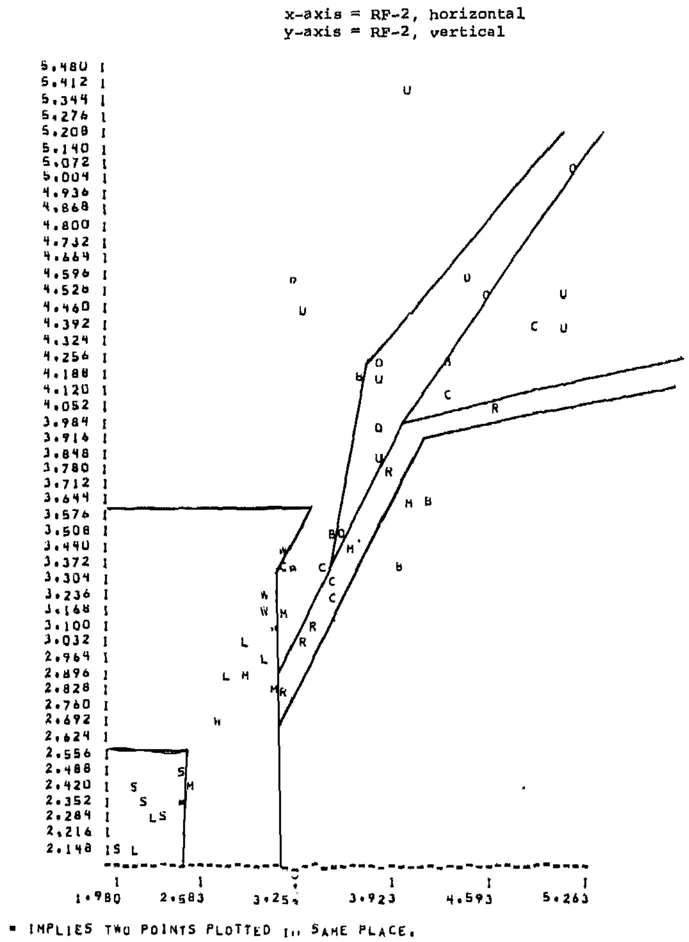

FIGURE 6

elimination process where the categories are divided into small groups. One example of this process is illustrated in Table 3. The category separations are indicated by spaces and the graph used is noted on the right.

These experimental results seem quite promising and indicate that further work should be done with different data sets. In addition, the run length functions should be used in conjunction with more sophisticated pattern classification techniques.

TABLE 1

Figure	X-axis	Y-axis
2	RF2—vertical	RF2—left diagonal
3	RF4—horizontal	RF4—right diagonal
4	RF5—left diagonal	RF5—horizontal
5	RF3—vertical	RF3—left diagonal
6	RF2—horizontal	RF2—vertical

TABLE 2

Category	Figure numbers
Swamp	2, 3, 4, 5
Lake	2, 3
Rail	2, 4, 5
Orchard	6
Scrub	2, 5
Suburb	2, 4

TABLE 3

Step no.	Category groups	Figure no.
1	S LW COUBRM	3
2	S LW C OUBM R	5
3	S L W C OUM B R	2
4	S L W C OU M B R	6
5	S L W C O U M B R	6

Two modifications to the run length texture analysis technique might be interesting to pursue. The first modification involves the computation of the run length matrices. Since diagonal neighbors are at a distance $u\sqrt{2}$, all diagonal run lengths should be mutliplied by $\sqrt{2}$. The second modification affects RF1. For the terrain samples, the values of RF1 were clustered. These values might tend to spread more if runs of length 1 were not used in the computation.

ACKNOWLEDGMENTS

The support of the Division of Computing, National Science Foundation, under Grant GJ-32258X, is gratefully acknowledged. The terrain samples were provided by Professor R. M. Haralick of the University of Kansas.

REFERENCES

1. J. K. Hawkins, Textural properties for pattern recognition, in *Picture Processing and Psychopictorics* (B. S. Lipkin and A. Rosenfeld, Eds), pp. 347–370, Academic Press, New York, 1970.
2. R. M. Haralick, K. Shanmugam, and I. Dinstein. "Textural features for image classification, *IEEE Trans. Systems, Man, and Cybernetics* SMC-3, No. 6, 1973, 610–621.
3. E. B. Troy, E. S. Deutsch, and A. Rosenfeld. Gray-level Manipulation Experiments for Texture Analysis, *IEEE Trans. Systems, Man, and Cybernetics* SMC-3, No. 1, 1973, 91–98.
4. M. M. Galloway. Texture Analysis Using Gray Level Run Lengths, Tech. Rept. 314, Computer Science Center, University of Maryland, July 1974.