A STUDY OF FEATURE EXTRACTION USING DIVERGENCE ANALYSIS OF TEXTURE FEATURES

W. Hallada, B. Bly and R. Boyd
Computer Sciences Corporation
Silver Spring, Maryland 20910, U.S.A.

S. Cox
NASA/GSFC, Code 902.1
Greenbelt, Maryland 20771, U.S.A.

ABSTRACT

This paper presents an empirical study of texture analysis for feature extraction and classification of high spatial resolution remotely sensed imagery (10 meters) in terms of specific land cover types. Little is known as to which texture features are important for separating specific land covers with a per-pixel classifier. The principal method examined is the use of spatial gray tone dependence (SGTD). The SGTD method reduces the gray levels within a moving window into a two-dimensional spatial gray tone dependence matrix which can be interpreted as a probability matrix of gray tone pairs. Haralick et al (1973) used a number of information theory measures to extract texture features from these matrices, including angular second moment (inertia), correlation, entropy, homogeneity, and energy. The derivation of the SGTD matrix is a function of: 1) the number of gray tones in an image; 2) the angle along which the frequency of SGTD is calculated; 3) the size of the moving window; and 4) the distance between gray tone pairs. In this study, the first three parameters were varied and tested on a 10 meter resolution panchromatic image of Maryville, Tennessee using the five SGTD measures. A transformed divergence measure was used to determine the statistical separability between four land cover categories--forest, new residential, old residential, and industrial--for each variation in texture parameters.

1.0 INTRODUCTION

With the successful launch of Landsat 4, remote sensing investigators will be receiving multispectral imagery of land areas at more than one spatial resolution, 30 meters from the Thematic Mapper (TM) and 82 meters from the Multispectral Scanner (MSS). In the future, multispectral linear array (MLA) technology will provide digital imagery of even higher spatial resolution, on the order of 10 to 15 meters in the visible and near infrared (NIR), 30 meters in the short wave infrared (SWIR), and 120 meters in the thermal infrared [1]. It is apparent that more innovative approaches to digitally extract information from mixed resolution systems need to be examined. In terms of spatial and spectral resolutions, the method and data used for extracting information will obviously depend upon the application, the level of computing advancements, and the associated costs and benefits obtained by using digital data [2].

One analysis technique commonly used is supervised or unsupervised per-pixel multispectral classification. A problem investigators have found with this technique is that as spatial resolution increases, classification accuracies can decrease for land covers of high spatial complexity, such as encountered in urban and tropical environments. Markham and Townshend [3] found that when a higher percentage of mixed pixels exist, classification accuracies decrease. Conversely, spectral heterogeneity of other land cover classes tended to be averaged out at lower spatial resolutions. This resulted in less spectral overlap with other land cover classes, which in turn resulted in higher classification accuracies. Latty [4] found similar results for forest cover classification.

Higher spatial resolution therefore compounds the classification problem if the spectral information is not used in context with the spatial information. The classifier must be able to characterize the spatial context of spectral reflectances for each land cover type. It becomes readily apparent that this information needs to be incorporated into the classification process to make the digital extraction of information from future satellite imagery successful.

A number of algorithms and approaches have been developed to include spatial information in the classification process. Townshend and Justice [5] provide a brief summary of popular methods. These include texture analysis described by Haralick [6], spectral/spatial context used by Tilton and Swain [7], and categorical/spatial context used by Wharton [8]. The purpose of this paper is to examine one particular method of texture analysis introduced by Haralick et al [9] to extract spatial features that are described by second order statistics.

1.1 Texture Analysis

One common approach used to characterize an image's spatial information is to extract features for classification which measure the spatial arrangement of gray tones within a neighborhood of a pixel. This feature extraction method is referred to as texture analysis and includes a multitude of possible features that have been developed to descibed image texture. Haralick [6] presents a complete literature review of texture analysis and Davis [10] presents some of the more recent developments. Conners and Harlow [11] investigated the mathematical and theoretical merits of various texture measures, whereas, Weszka et al [12] conducted an empirical comparison of various texture measures. Unfortunately, as noted by Townshend and Justice [5], more effort has been expended on the derivation of new texture measures than on evaluating the relative merits of each method for remotely sensed data. The use of texture analysis has been hindered with current satellite imagery because it effectively coarsens the spatial resolution. This introduces edge effects at the boundaries of land covers comprised of different gray tones or textures. Calculating a texture measure for a 5-by-5 window passed over a Landsat image coarsens the resolution to 400 meters. Future MLA resolutions should mitigate such effects.

Of the numerous texture measures available, the spatial gray tone dependence (SGTD) method has been used frequently by remote sensing investigators including Haralick et al [9], Haralick [13], Jensen [14], Jensen and Toll [15], Schowengerdt [16], and Weszka et al [12]. SGTD represents, both conceptually and computationally, an approach of greater breadth and complexity to texture extraction than such first order statistics as mean and standard deviation. The SGTD method transforms the gray values within a neighborhood of each pixel into a two-dimensional gray value co-occurrence matrix. This matrix $P(i,j|\alpha,d)$ describes the frequency of occurrence of gray value pairs i, j separated by distance d, and angle α, and therefore can be interpreted as a probability matrix of gray value pairs. Haralick et al [9] was the first to introduce a number of measures based on information theory to describe such matrices. Cox and Rose [17] developed computationally efficient software for calculating SGTD textures within the Interactive Digital Image Manipulation System (IDIMS). Measures implemented to date include: inertia (angular second moment difference); homogeneity (angular second moment inverse difference); correlation (covariance of neighboring pixels); entropy (average uncertainty of gray values); and, energy (angular second moment). These measures are mathematically summarized in Table 1, and are described in Table 2.

TABLE 1

- INERTIA (Angular Second Moment Difference)

$$INT = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} (1-j)^2 \frac{P(i,j|\alpha,d)}{\#R}$$

- ENERGY (Angular Second Moment)

$$ENG = \sum_{i=1}^{Ng} \sum_{j-1}^{Ng} \left[\frac{P(i,j|\alpha,d)}{\#R}\right]^2$$

- ENTROPY (Average Uncertainty of $P(i,j|\alpha,d)$)

$$ENT = -\sum_{i=1}^{Ng} \sum_{i=j}^{Ng} \frac{P(i,j|\alpha,d)}{\#R} \text{Log} \left(\frac{P(i,j|\alpha,d)}{\#R}\right)$$

- HOMOGENEITY (Angular Second Moment Inverse Difference)

$$HOM = \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} \frac{1}{1 + (i+j)^2} \frac{P(i,j|\alpha,d)}{\#R}$$

- CORRELATION (Covariance of Neighboring Points)

$$COR = \frac{\sum_{i=1}^{Ng} \sum_{j=1}^{Ng} (ij) \frac{P(i,j|\alpha,d) - \mu_x \mu_y}{\#R}}{\sigma_x \sigma_y}$$

WHERE #R = Number of Neighboring Cells
N_g = Number of Gray Tones, and μ_x, μ_y, σ_x, σ_y are the means and standard deviations of the marginal distributions associated with $P(i,j|\alpha,d)/\#R$.

TABLE 2

- INERTIA
  - Measures tendency to concentrate probability away from the main diagonal of the co-occurrence matrix.
  - Related to gray value variance.
  - Inversely proportional to image coarseness, or contrast.
  - Lower bound when texture is entirely monotone.
- HOMOGENEITY
  - Measures the similarity of neighboring pixels.
  - Flat textures will give higher values.
  - Upper bound when all probability lies on the main diagonal of the co-occurrence matrix.
- CORRELATION
  - Measures the covariance of neighboring pixels.
  - Zero when all pixels are independent.
  - Natural scenes tend to have lower values.
  - Has the largest values for periodic patterns.
- ENTROPY
  - Measures the average uncertainty of gray values pairs.
  - Upper bound when all probabilities are equal.
  - Lower bound when one gray tone pair has a probability of 1.
  - Invariant to monotonic gray tone transformations.
- ENERGY
  - Measures the average certainty of occurrence of gray value pairs.
  - Lower bound when all probabilities are equal.
  - Upper bound when only one probability appears.
  - Homogeneous areas have higher energy.
  - Invariant to monotonic gray value transformations.

In summary, the derivation of the SGTD matrix is a function of the following parameters:

1. The number of gray levels within an image. The computation of the texture feature is related to the square of the number of gray levels.

2. The angle along which the frequency of occurrence is derived. For example, there are four independent angles for a distance of one, and eight for a

distance of two resulting in four and eight independent features for each image. Haralick et al [9] suggested using the average and range (which are invariant under rotation) as inputs to the classifier.

3. The size of the moving window. Small window sizes will not adequately sample the SGTD probabilities of land cover classes [18]. Conversely larger window sizes will degrade the resolution of remotely sensed imagery.

4. The distance between pixels in tabulating the co-occurrence matrix. Haralick [6] argued that the co-occurrence matrix for a single distance contains most of the significant texture information.

An empirical investigation of the effects of quantization level, orientation, and window size upon classification accuracy of remotely sensed imagery does not, to our knowledge, exist in the literature. Little is known about the effectiveness of texture analysis in terms of sensor spatial resolution and the spatial frequencies of land cover on the ground. In addition to the above parameters, approximately two dozen dependent SGTD features can be used [6]. Compression techniques using eigenvector analysis were proposed by Tou and Chang [19] to reduce this large dimensionality for features comprised of SGTD measures from different angles and distances.

2.0 STUDY SITE AND TEST DATA

For the texture investigations, a test site containing a mixture of urban, forest, and agricultural land covers was chosen in order to provide a variety of textures to study. The digital imagery was acquired using a Daedalus DS-1260 MSS flown on April 7, 1977 over Maryville, Tennessee. Flown from an altitude of 3,000 meters, the instantaneous field of view (IFOV) at nadir was 8.25 meters. This data set was one of many processed by Geospectra Corporation of Ann Arbor, Michigan under contract to the National Oceanic and Atmospheric Administration to provide multispectral imagery of diverse land cover types. In processing the data, Geospectra Corporation resampled the original scanner data to 10 meter resolution and rectified it to a Universal Transverse Mercator projection. Additional processing included interband averaging to simulate Thematic Mapper bands, and contrast stretching. The contrast stretching, interband averaging, resampling, and degradation methods made the utility of the data questionable [20]. For future studies it is suggested that the methods reported herein be attempted with data more representative of future MLA satellite instruments.

The study site, Maryville, Tennessee is a small city of 14,000 people located 10 miles south of Knoxville. The entire test site includes a range of residential densities, commercial and industrial areas including infrastructure such as roads and airports, forested areas and agricultural fields. The entire test image covers an area of approximately 5 square kilometers, with the street pattern oriented at 45 degrees off the image line and sample axes. Visual interpretation of this data, clearly reveals that applications of remote sensing for urban studies would readily benefit from a 10 meter MLA instrument.

3.0 METHODS

To reduce the number of features and preserve as much spatial information as possible, a panchromatic image was synthesized from the green (.55-.60 um), red (.6-.69 um) and near-infrared (.8-1.1 um) bands. The norm, or the square root of the sum of the three squared gray values was used to simulate a panchromatic image. The correlation of the panchromatic band with the green, red, and near-infrared bands was .89, .93, and .82, respectively. The resulting image is shown in Figure 1.

3.1 Test Site Selection

From this image four training sites were chosen to study the effect of window size, quantization, orientation, and SGTD measure upon classification accuracy. Each test site was selected based on land cover, visual texture, and size. The four sites were: 1) mature deciduous forest; 2) old residential composed of mature deciduous trees, old homes and narrow paved roads; 3) new residential composed of large lots, larger ranch style homes, wider roads and few trees; and, 4) an industrial site with concrete parking lots and large buildings with linear shadows. Roads in both residential sites were oriented at 45 degrees and the buildings in the industrial site horizontally. The four unique complex land covers, shown in Figure 1, provided a good basis for comparing texture features.

Each test site consisted of a 40-by-40 pixel block. After applying the five texture measures on a pixel by pixel basis, statistics were calculated for a 20-by-20 pixel training block centered within each test site to eliminate any harmful effects from edges between sites, as well as to provide a large sample of 400 pixels per land cover.

3.2 Feature Extraction

The effects of quantization, window size, and orientation angle were extensively tested using the inertia texture measure (Table 1). Only the window size parameter was varied for the other four texture measures of energy, entropy, correlation, and homogeneity. The quantization level was tested by requantizing the gray levels from 256 levels to 4, 8, 16, 32, 64 and 128 gray levels using a simple linear mapping. Spatial gray tone inertia features were then generated using sliding windows of 3-by-3 (30 meter textures) to 13-by-13 (130 meter textures), in increments of two, for the four possible orientations (0, 45, 90 and 135 degrees) at a distance of one. A total of 168 texture features were therefore created and evaluated; that is, 4 orientations, 6 window sizes, and 7 quantization levels. The means and covariances of the four orientations combined with the original gray tone image were calculated for each quantization level, window size, and training site.

3.3 Transformed Divergence Analysis

Rather than perform an actual classification for various combinations of texture features followed by an accuracy test using test and training sites, a statistical measure of separability was employed as a predictor of classifier performance. Once the statistics were calculated for each feature set combination, a transformed divergence measure was used to determine the interclass separability of the four land covers.

Divergence [21] between class pairs i and j is defined as:

$$D_{ij} = 1/2 \text{ tr} \left[(\Sigma_i - \Sigma_j) (\Sigma_j^{-1} - \Sigma_i^{-1}) \right] + 1/2 \text{ tr} \left[(\Sigma_i^{-1} + \Sigma_j^{-1}) (M_i - M_j) (M_i - M_j)^T \right]$$

WHERE Σ = Class covariance matrix

M = Vector of class means
tr = trace (sum of the diagonal elements).

Because divergence increases without bound as statistical separability between classes increases, Swain and Davis [21] defined a saturation transform which provides a measure more closely corresponding to percent correct classification. The transformed divergence expression is:

$$TD_{ij} = 2{,}000 \left[1 - \exp (-D_{ij}/8) \right].$$

This measure has a saturating behavior, that is, percent correct classification saturates at 100 percent when a certain level at statistical separability is reached (TD = 2,000).

There are some disadvantages in using transformed divergence as a measure of statistical separability between class pairs. For example, two class densities having equal mean vectors but non-equivalent covariance matrices may result in a transformed divergence of zero [22]. Furthermore, there is no estimate for a lower confidence limit for the relation between transformed divergence and percent correct classification. In lieu of alternative measures, transformed divergence is very efficient computationally, and affords a relative measure of performance without doing an actual classification.

4.0 RESULTS

The average transformed divergences (TD) of all land cover pairs are plotted in Figure 2 for each window size of inertia calculated from data with 128 gray levels. As the window size increased, the average TD value increased. Combining the four orientations into a single normalized measure significantly reduced the average TD. The addition of the gray tone image increased separability but not enough for acceptable classification accuracy. The increases in average TD of the four orientations behaved in a logarithmic fashion and began to level off at a window size of 11 pixels (110 meters).

Figure 3 plots the TD values using four orientations of inertia for each land cover pair except for those having forest. The TD between forest and all other land covers was 2000. Apparent length of the lines connecting different window sizes in Figure 3 is proportional to the added separability resulting from an increase in spatial information. Four orientations of inertia should, therefore, provide features which may be used to classify forest with 100 percent accuracy at any window size. As expected, the separability between the two residential classes was the lowest for all window sizes. A larger window size may be needed to improve the separability between these two categories.

The effect of gray level quantization upon TD is shown in Table 3 for the two residental categories. A decrease in separability accompanied by a decrease in gray level quantization does not occur until approximately level 16. At this level the separability between the two land covers decreases somewhat. A larger window size did compensate for this decrease.

TABLE 3

Transformed divergence between new and old residential for changes in window size and gray level quantization.

Quantization Level	Transformed Divergence Values					
	3x3	5x5	7x7	9x9	11x11	13x13
4	184	413	668	1120	1430	1627
8	251	558	874	1325	1584	1699
16	238	519	822	1266	1638	1797
32	265	563	875	1319	1667	1800
64	256	553	869	1319	1657	1796
128	256	549	867	1323	1672	1801
256	258	556	874	1327	1675	1806

Quantization level may affect both the feature extraction and classification process, and hence affect accuracies. As noted earlier, reducing the number of gray tones in the input image will decrease the size of the SGTD matrix as well as reduce computation time. Furthermore, most maximum likelihood classification software in image processing systems are written to classify byte data with 256 gray levels. The result of a texture transform is a real number which must be scaled and quantized to fit within 256 gray levels.

Table 4 shows the result of requantizing the four orientations of inertia for new and old residential. Similarly, there was a reduction in separability due to the requantization of the texture measure, but the reduction does not seem significant.

TABLE 4

Transformed Divergence
Between Old and New Residential.

Window Size	256 Levels	Floating Point
7x7	736	875
13x13	1710	1800

Additional insight into the effect of orientation on texture feature extraction was gained through plotting transformed divergences for subsets of 2 and 3 orientations, as shown in Figure 4. From this figure, it is evident that all four orientations were important for separating various urban land covers. No subset of 2 or 3 orientations provided adequate separation between all land covers, although larger window size may compensate for the loss of orientation features.

The difference in separability due to the type of SGTD measure is shown in Figure 5. Homogeneity was better at separating the industrial and residential land covers, whereas inertia provided the best feature for separating forest from all other land cover categories. The average transformed divergence plotted in Figure 6 indicates that inertia had the best overall separability performance for separating the four land cover types, and correlation was the worst.

5.0 CONCLUSIONS

The results obtained in this empirical study demonstrated that quantization level, window size, and orientation are very important parameters to consider when using the SGTD method for extracting texture features from high spatial resolution remotely sensed imagery. Although transformed divergence did not provide a perfect measure of classification accuracy; it did provide a robust method to evaluate texture features.

In summary, the following results were found:

1) As window size increased, class separability increased logrithmically. Separability between certain land covers was maximized at smaller window sizes depending upon the SGTD measure.

2) Class separability was very sensitive to SGTD orientation. A subset of orientations as well as the norm of the four orientations proved inadequate for separating the four land covers.

3) Class separabilities did not begin to decrease with decreasing gray levels until 16 gray levels. At 16 or 8 gray levels larger window sizes were needed to preserve separability.

4) Rescaling texture features from a 32-bit real number into 256 gray levels reduced class separability; however, this did not seem as important a factor as window size and orientation.

5) Of the five SGTD measures tested, inertia had the best performance for separating the few land covers investigated herein.

6.0 RECOMMENDATIONS

Quantitative criteria should be developed to determine the spatial resolutions optimal for using texture analysis for specific applications, e.g. urban remote sensing. Work similar to that reported herein should be attempted with remotely sensed data of various spatial and spectral resolutions, and to include further empirical comparisons of other texture features. These feature combinations should encompass first order statistics [6] and the recently developed texture energy measures [10]. Additionally, it is recommended that algorithms, which process image data spatially be implemented on parallel processors such as the massively parallel processor [23].

Furthermore, it is suggested that recent developments in image texture analysis, as reported in the pattern recognition literature, be attempted with remotely sensed imagery. Such recent developments are: the use of the co-occurrence matrix directly in a classification algorithm [18, 24]; the use of segmentation techniques to partition remotely sensed imagery into unique texture regions, such that boundaries between regions are correctly represented [10, 25]; and, research into multispectral texture models and classification [26].

7.0 REFERENCES

1. Mika, A.M., 1982. Design tradeoffs for a multispectral linear array (MLA) instrument. In the Multispectral Imagery Science Working Groups: Final Report, Vol. 3, Contributed Papers. S. Cox, Ed. September 1982. NASA Conference Publication in Press.

2. Welch, R., 1982. Spatial resolution requirements for urban studies. International Journal of Remote Sensing. 3: 139-146.

3. Markham, B.L. and Townshend, J.R.G., 1981. Land cover classification accuracy as a function of sensor spatial resolution. Proceedings of the 15th International Symposium on Remote Sensing of Environment, Ann Arbor, Michigan (in press).

4. Latty, R.S., 1981. Computer based forest cover classification using multispectral scanner data of different spatial resolutions. LARS Technical Report 052081. Laboratory for Applications of Remote Sensing, West Lafayette, Indiana, U.S.A.

5. Townshend, J. and Justice, C., 1981. Information extraction from remotely sensed data. A user view. International Journal of Remote Sensing. 2: 313-329.

6. Haralick R.M., 1979. Statistical and structural approaches to texture. Proceedings of the IEEE. 67: 786-804.

7. Tilton, J.C. and Swain, P.H., 1981. Incorporating spatial context into spatial classification of multidimensional image data. LARS Technical Report 072981. Laboratory for Applications of Remote Sensing, West Lafayette, Indiana, U.S.A.

8. Wharton, S.W., 1982. A contextual classification method for recognizing land use patterns in high resolution remotely sensed data. Pattern Recognition. 15:317-324.

9. Haralick, R.M., Shanmugam, K., and Dinstein, I., 1973. Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics. SMC-3:610-621.

10. Davis, L.S., 1982. Image texture analysis: Recent developments. In Proceedings of 1982 IEEE Conference on Pattern Recognition and Image Processing, Las Vegas, Nevada, pp. 214-217.

11. Conners, R.W., and Harlow, C.A., 1980. A theoretical comparison of texture algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence. PAMI-2:204-222.

12. Weszka, J.S., Dyer, C.R., and Rosenfeld, A., 1976. A comparative study of texture measures for terrain classification. IEEE Transactions on Systems,Man and Cybernetics. SMC-6:269-285.

13. Haralick, R.M., 1974. Combined spectral and spatial processing of ERTS imagery data. Journal of Remote Sensing of the Environment. 3:3-13.

14. Jensen, J.R., 1979. Spectral and textural features to classify elusive land cover at the urban fringe. The Professional Geographer. 31:400-409.

15. Jensen, J.R. and Toll, D.L. 1982. Detecting residential land-use developments at the urban fringe. Photogrammetric Engineering and Remote Sensing. 48:629-643.

16. Schowengerdt, R.A., 1981. Texture feature extraction. Final Report, USGS Grant #14-08-0001-G-664. Applied Remote Sensing Program, University of Arizona, Tucson, Arizona, U.S.A.

17. Cox, S. and Rose, J., 1982. Texture functions in image analysis, computationally efficient solution. In press.

18. Vickers, A.L. and Modestino, J.W., 1982. A maximum likelihood approach to texture classification. IEEE Transactions on Pattern Analysis and Machine Intelligence. PAMI-4:61-68.

19. Tou, J.T. and Chang, Y.S., 1977. Picture understanding by machine via textural feature extraction. In Proceedings 1977 IEEE Conference on Pattern Recognition and Image Processing. (Troy, New York), pp. 392-99.

20. Communication with EROS Data Center, May, 1982.

21. Swain, P.H. and Davis, S.M., 1978. Remote Sensing: The Quantitative Approach, New York: McGraw-Hill.

22. Swain, P.H., Robertson, T.V., Wacker, A.G., 1971. Comparison of the divergence and B-distance in feature selection. LARS Information Note 200871. Laboratory for Applications of Remote Sensing, West Lafayette, Indiana, U.S.A.

23. Strong, J.P., Schaeffer, D.H., Fischer, J.R., Wallgren, F.A., and Bracken, P.A., 1978. The massively parallel processor and its applications. Proceedings of the 12th International Symposium on Remote Sensing of Environment, ERIM, University of Michigan, Ann Arbor, pp. 37-45.

24. Zucker, S.W. and Terzopoulos, D., 1982. Finding structure in co-occurrence matrices for texture analysis. In Image Modeling. A. Rosenfeld, ed. New York: Academic Press, pp. 423-445.

25. Davis, L.S. and Mitichie, A., 1982. MITES: A model-driven, iterative texture segmentation algorithm. Computer Graphics and Image Processing. 19:95-110.

26. Rosenfeld, A., Wang, C., and Wu, A.Y. Multispectral texture. IEEE Transactions on Systems, Man, and Cybernetics. SMC-12: 79-84.

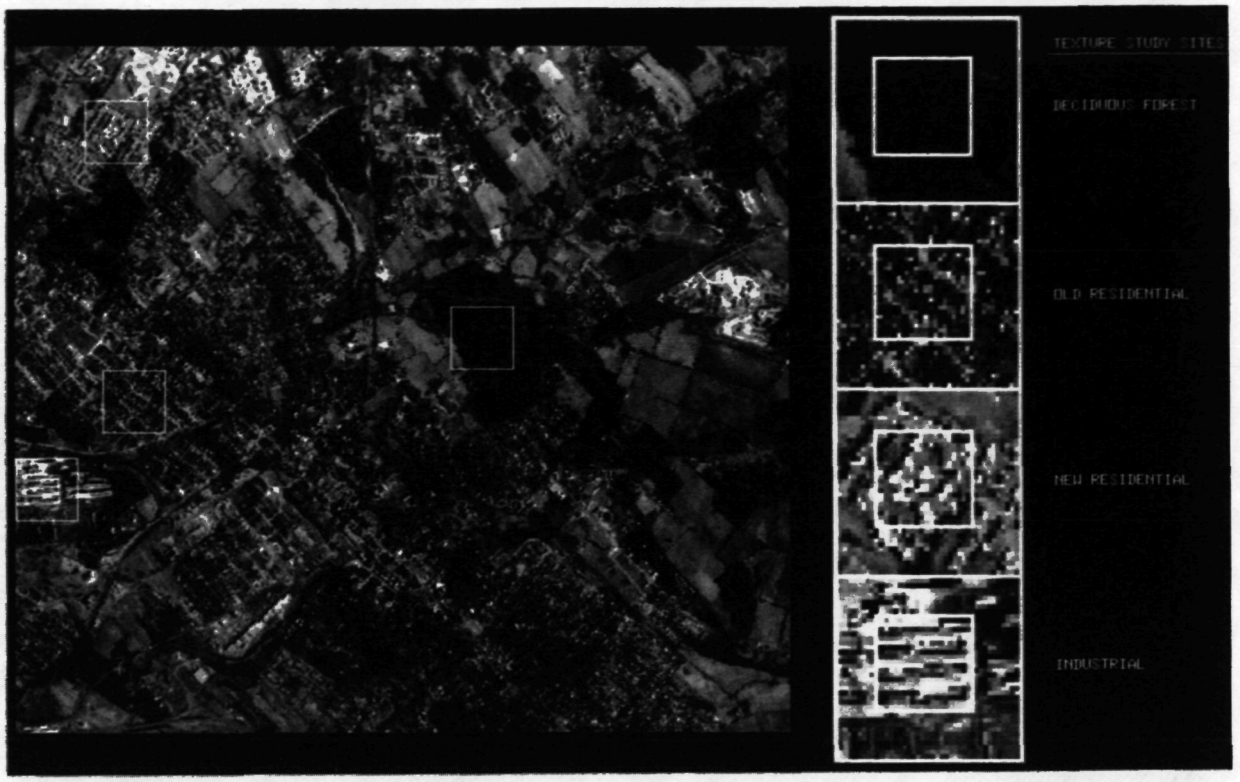

Figure 1. Synthesized panchromatic image and texture study sites.

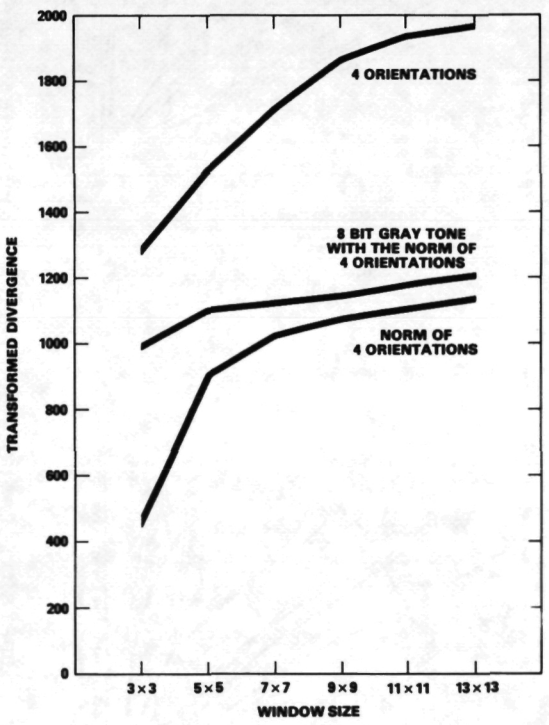

Figure 2. Average separability using 4 orientations of inertia, versus the norm and norm plus panchromatic image.

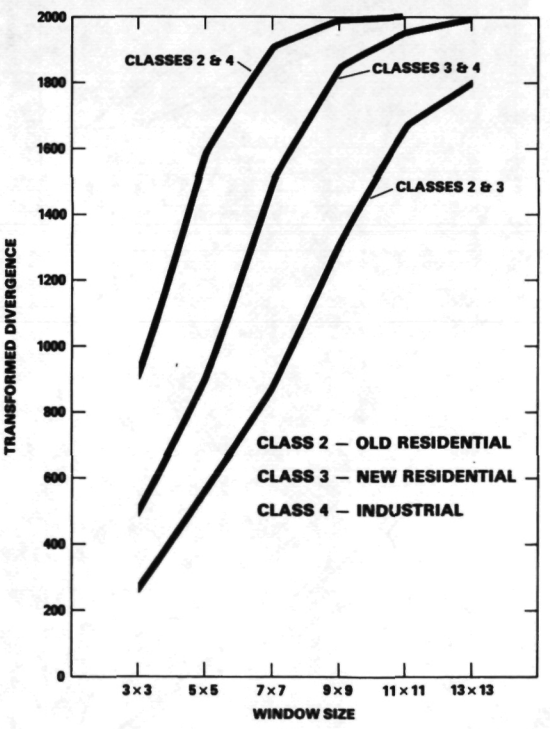

Figure 3. Separability between land covers using 4 orientations of inertia.

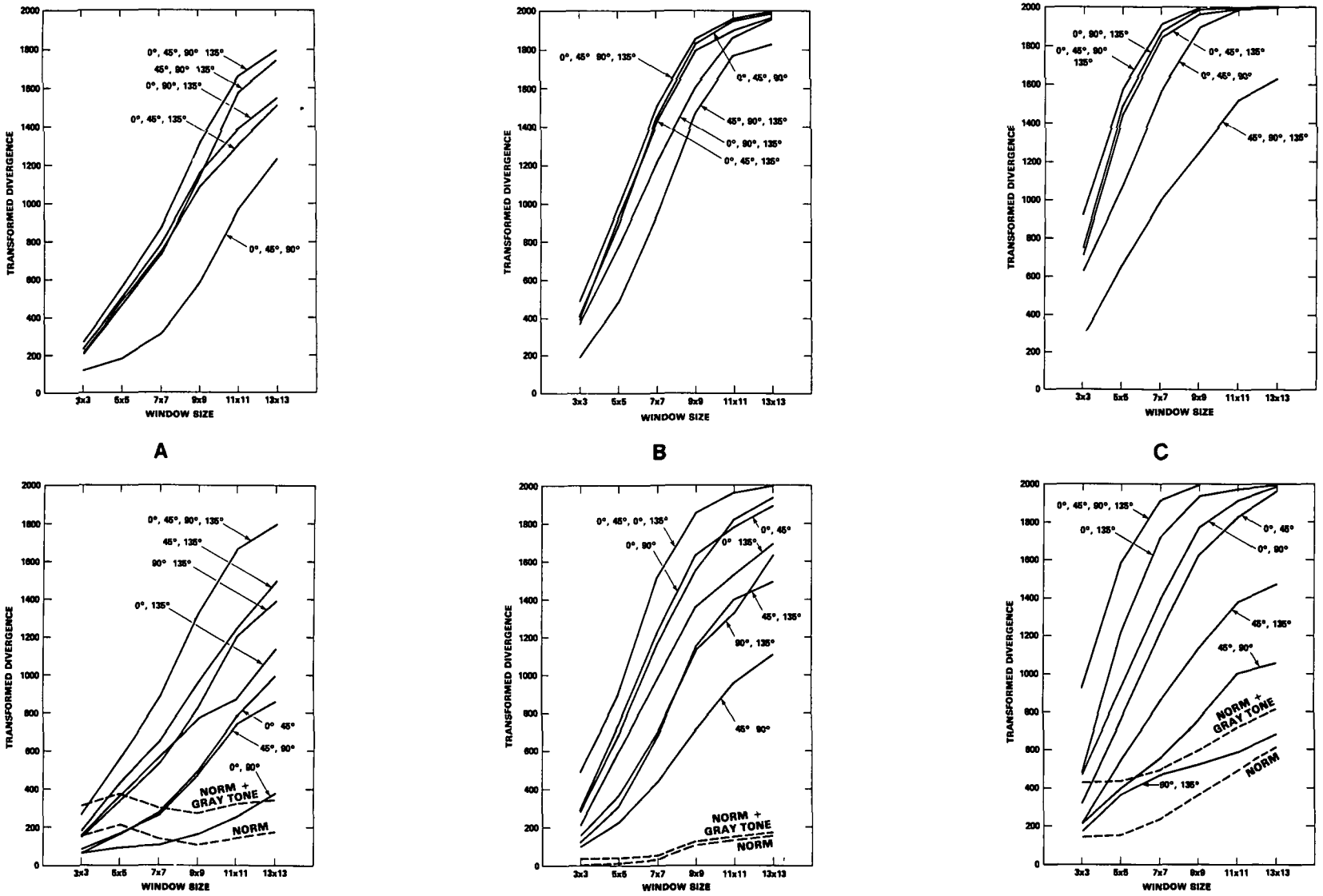

Figure 4. Separability between land covers using spatial gray tone inertia, a) new and old residential, b) new residential and industrial, and c) old residential and industrial.

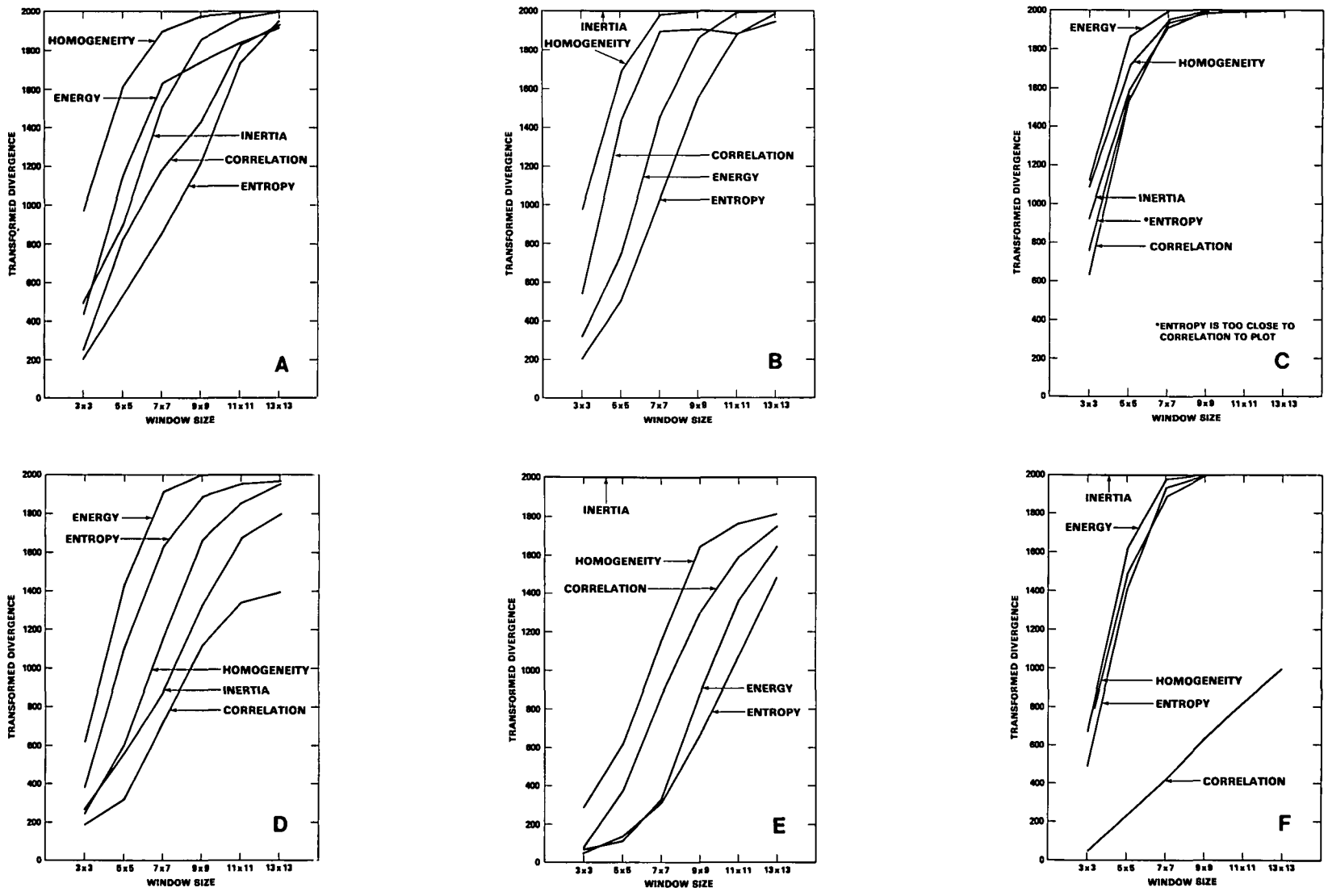

Figure 5. Separability between a) new residential and industrial, b) industrial and forest, c) old residential and industrial, d) new and old residential, e) new residential and forest, and f) old residential and forest.

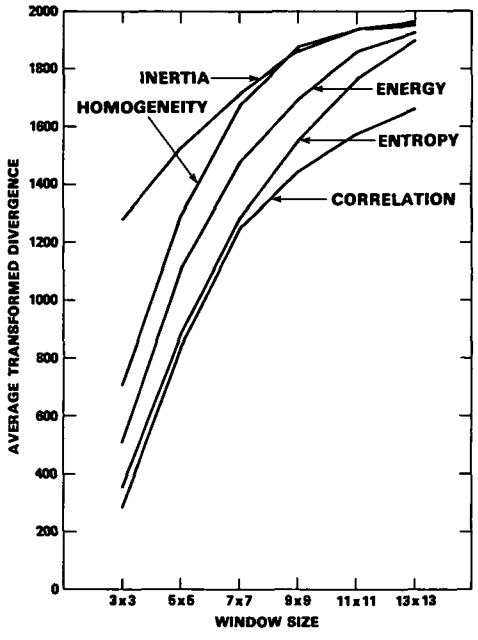

Figure 6. Average interclass separability between land covers using 4 orientations of SGTD measures.