SPECTRAL AND TEXTURAL PROCESSING OF ERTS IMAGERY

R. M. Haralick and R. Bosley, The University of Kansas Center for Research, Inc., Lawrence, Kansas 66044

ABSTRACT

A procedure is developed to simultaneously extract textural features from all bands of ERTS multi-spectral scanner imagery for automatic analysis. Multi-images lead to excessively large grey tone N-tuple co-occurrence matrices; therefore, neighboring grey tone N-tuple differences are measured and an ellipsoidally symmetric functional form is assumed for the co-occurrence distribution of multi-image greytone N-tuple differences. On the basis of past data the ellipsoidally symmetric approximation is shown to be reasonable. Initial evaluation of the procedure is encouraging.

1. INTRODUCTION

A procedure is being developed to extract textural features for automatic analysis of ERTS multi-spectral scanner imagery. Previous work (Haralick, 1973; Haralick, 1972) indicates that useful textural features can be computed from co-occurrence matrices for grey tones in specific spatial relationships on an image. The performance of the land use classification algorithm using these textural features from only one band is encouraging; up to 75 per cent of the images were correctly classified (Haralick, 1973). Since texture features and spectral features of ERTS multi-images provide different kinds of information, a significant increase in identification accuracy will occur when both features are used together.

Multi-images lead to excessive amounts of storage for the grey tone N-tuple co-occurrence matrices. Therefore, to solve the storage problem we measure grey tone N-tuple differences instead of grey tone N-tuples and assume an ellipsoidally symmetric function form for the co-occurrence distribution of multi-image grey tone N-tuple differences. We expect that the estimated parameters of the ellipsoidally symmetric distribution will lead to textural features which will distinguish between texturally distinct categories on ERTS MSS images over Kansas.
II. TEXTURE

Texture and tone are two fundamental pattern elements used in the interpretation of image data. The concept of tone is concerned with the whiteness, greyness, or blackness of resolution cells of the image. The concept of texture is concerned with the spatial distribution of the grey tones. Tone is based upon the varying shades of grey of the resolution cells in the image, while texture is based upon the spatial distribution of grey tones. However, texture and tone are not independent concepts but are intrinsically related to one another. Although either property can dominate the other depending upon the image context, texture and tone are always present.

When one attempts to objectively use tone and texture pattern elements, the texture-tone concept must be explicitly defined. This can be visualized as follows. When a small area patch of an image has little variation of features of discrete grey tone, then that area is dominated by tonal properties. As the number of distinguishable features of discrete grey tone increases within the patch, then the texture properties will dominate. The size of the small area patch, the relative sizes of the discrete features and the number of distinguishable discrete features are all crucial in this distinction. When the size of the small area patch is reduced to one resolution cell, the only property present is tone. When there is no spatial pattern in the tonal features and the grey tone variation between features is wide, a fine texture results. And as the spatial pattern becomes more defined using more and more resolution cells, then a coarser texture results.

Texture can be termed as being fine, coarse, smooth, rippled, molled, irregular, or lineated. Texture is a property of nearly all surfaces, the grain of wood, the weave of fabric, the pattern of crops in a field, etc. Although texture is quite easy for humans to recognize and describe, it is quite subjective by its nature and is extremely difficult to precisely define and analyze by digital computers. Since the texture of images contains important information for discrimination purposes, textural features could be very useful.
III. REVIEW OF PAST WORK ON TEXTURE

To date there has been at least six different approaches to the problems of measuring and characterizing texture of images: autocorrelation functions, optical transforms, digital transforms, edgeness, structural elements, and spatial grey tone co-occurrence probabilities. The first three approaches all measure spatial frequency either directly or indirectly. Spatial frequency is related to texture because fine textures are rich in high frequencies while coarse textures are rich in low frequencies.

One alternative approach to viewing texture as spatial frequency distribution is to view texture as the amount of edgeness per unit area. Fine textures have a high number of edges per unit area whereas coarse textures have a small number of edges per unit area.

The structural element approach uses a matching procedure to detect the spatial regularity of shapes called structural elements in a binary image. When the structural elements themselves are single resolution cells, the information provided by this approach is the autocorrelation function of the binary image. By using larger and more complex shapes, a more generalized autocorrelation can be computed.

The grey tone spatial dependence approach characterizes texture by the spatial distribution of its grey tones. In coarse textures the distribution changes only slightly with distance, but for fine textures it changes rapidly with distance.

Because of our familiarity with the concepts of spatial frequency and edgeness, these approaches to texture characterizations are readily employed. However, an inherent problem exists with these approaches in regard to grey tone calibration of the image and they are not invariant under even a linear grey tone translation. And the price paid for invariance by compensating with quantization is a loss of grey tone precision in the quantized image.

The power of the structural element approach is that it emphasizes the shape aspects of the discrete tonal features. Weakness of this approach lies in that it can only do so for binary images.

The power of the spatial grey tone co-occurrence approach lies in characterizing the spatial inter-relationships of the grey tones in a texture pattern in such a way that is invariant under monotonic grey tone transformations. Weakness of the approach lies in failure to capture the shape aspects of the discrete tonal features.
IV. TEXTURAL FEATURES

The above description of texture is an idealization of what actually occurs, a gross simplification. Discrete tonal features are actually quite subjective in that they do not necessarily stand out as entities by themselves. Therefore, the texture analysis presented here is concerned with more general or macroscopic concepts rather than discrete tonal features. The procedure developed by Haralick (Haralick, 1972) for obtaining the textural features of an image is based on the assumption that the texture information on an image \( I \) is contained in the overall spatial relationship which the greytones in the image \( I \) have to one another. More specifically, we assume that this texture information is adequately specified by a set of spatial grey tone dependence matrices, which are computed for various angular relationships and distances between neighboring resolution cell pairs on the image. All of the textural features are then derived from these angular nearest neighbor spatial grey tone dependence matrices.

IV.1 Spatial Grey Tone Dependence Matrices

Let \( G = \{0, 1, \ldots, N_g\} \) be the set of possible grey tones that each resolution cell can take on after image normalization by equal probability quantizing to \( N_g \) levels. It can be shown that this quantization guarantees that images which are a monotonic transformation of one another, such as lighter or darker images due to variations in film, lighting, or development, will produce the same results. Let \( N_x \) be the number of resolution cells in the horizontal direction and \( N_y \) the number of resolution cells in the vertical direction in the image to be analyzed so that \( L_x = \{1, 2, \ldots, N_x\} \) and \( L_y = \{1, 2, \ldots, N_y\} \) are the horizontal and vertical spatial domains. Then \( L_y \times L_x \) will be the set of resolution cells of the image.

And the image \( I \) can be represented as a function which assigns some grey tone in \( G \) to each resolution cell or pair of coordinates in \( L_y \times L_x \); \( I:L_y \times L_x \rightarrow G \).

Essential to our conceptual framework of texture are four closely related measures called angular nearest neighbor grey tone spatial dependence matrices. The concept of angular nearest neighbor for a resolution cell is the adjacent resolution cell for a given angle, as shown in Figure 1.
FIGURE 1. Eight nearest neighbor resolution cells of cell \(1^\circ\).
Resolution cells 1 and 5 are the 0-degree nearest neighbors to resolution cell \(1^\circ\), resolution cells 2 and 6 are the 135-degree nearest neighbors, etc.
Note that this information is purely spatial, having nothing to do with grey tone values.

We assume that the texture information in our image \(I\) is contained in the overall or "average" spatial relationship which the grey tones in image \(I\) have to one another. Specifically, we shall assume that this information is adequately specified by the matrix of relative frequencies \(P_{ij}\) with which two neighboring resolution cells separated by distance \(d\) occur on the image, one with grey tone \(i\) and the other with grey tone \(j\). These matrices of spatial grey tone dependence frequencies are a function of the angular relationship between the neighboring resolution cells as well as a function of the distance between them. Figure 2 illustrates the set of all horizontal neighboring resolution cells separated by distance 1. This set along with the image grey tones would be used to calculate a distance \(i\) horizontal spatial grey tone dependence matrix. Formally, for angles quantized to \(45^\circ\) intervals the unnormalized frequencies are defined by:

\[
P(i, j; d, 0^\circ) = \# \{(k, l), (m, n) \in (L \times L) \times (L \times L) | k-m=d, \ |l-n|=d, \ 0 \leq k \leq l \leq 1, \ 0 \leq m \leq n \leq 1, \ k \leq m \leq l \leq n \}
\]

\[
P(i, j; 45^\circ) = \# \{(k, l), (m, n) \in (L \times L) \times (L \times L) | k-m=d, \ 0 \leq k \leq l \leq 1, \ 0 \leq m \leq n \leq 1, \ k \leq m \leq l \leq n \}
\]

\[
P(i, j; 90^\circ) = \# \{(k, l), (m, n) \in (L \times L) \times (L \times L) | k-m=d, \ 0 \leq k \leq l \leq 1, \ 0 \leq m \leq n \leq 1, \ k \leq m \leq l \leq n \}
\]

\[
P(i, j; 135^\circ) = \# \{(k, l), (m, n) \in (L \times L) \times (L \times L) | k-m=d, \ 0 \leq k \leq l \leq 1, \ 0 \leq m \leq n \leq 1, \ k \leq m \leq l \leq n \}
\]

Note that these matrices are symmetric; \(P(i, j; d, a) = P(j, i; d, a)\). The distance metric \(\rho\) implicit in the above equations can be explicitly defined by

\[
\rho((k, l), (m, n)) = \max \{|k-m|, |l-n|\}
\]
\( R_H = \{ (\{k, 1\}, (m, n)) \in (L_y \times L_x) \times (L_y \times L_x) | k - m = 0, |l-n| = 1 \} \)

\[
\begin{array}{cccc}
(1, 1) & (1, 2) & (1, 3) & (1, 4) \\
(2, 1) & (2, 2) & (2, 3) & (2, 4) \\
(3, 1) & (3, 2) & (3, 3) & (3, 4) \\
(4, 1) & (4, 2) & (4, 3) & (4, 4) \\
\end{array}
\]

\( L_y = \{1, 2, 3, 4\} \)
\( L_x = \{1, 2, 3, 4\} \)

**FIGURE 2.** Illustrates the set of all Distance 1 Horizontal Neighboring Resolution Cells on a 4 x 4 Image.
For an example of the four distance 1 grey tone spatial dependence matrices, consider Figure 3. Figure 3-a represents a 4x4 image with four grey tones, ranging from 0 to 3. Figure 3-b shows the general form for any grey tone spatial dependence matrix. For example, the element in the (2, 1)-st position of the distance 1 horizontal $P_H$ matrix is the total number of times two grey tones of value 2 and 1 occurred horizontally adjacent to each other. To determine this number, we count the number of pairs of resolution cells in $R_H$ such that the first resolution cell of the pair has grey tone 2 and the second resolution cell of the pair has grey tone 1. Figure 3-c through 3-f shows all four distance 1 grey tone spatial dependence matrices.

IV.2 Textural Features for Multi-Images

Results of previous work in texture using the spatial grey tone dependence matrices as the basis from which all textural features are extracted has been very encouraging (Haralick, 1973). The computational aspects of the procedure are also notable. The number of operations required to process an image using the spatial grey tone dependence matrices is directly proportional to the number of resolution cells, N, present in an image. In comparison, the number of operations needed to use Fourier or Hadamard transforms to extract texture information are of the order of $N \log N$. And, to compute the entries in the spatial grey tone dependence matrices, one needs to keep only two lines of image data in core at a time, keeping storage requirements to a minimum.

Even with these advantages, however, the extraction of texture information from multiimages, as in the case of ERTS MSS data, forces a new approach to the measurement of grey tone N-tuple co-occurrences. The use of the spatial dependence matrices requires that they be stored in the computer. For multi-images containing grey tone N-tuples, we have many more possible combinations of neighboring N-tuples to count and as a result, the dependence matrices will be very large. For example, for four MSS bands in which each grey tone can range through 64 levels, each matrix would have $64^4 \times 64^4$ elements. Even using the symmetry of the matrices to reduce the number of entries does not help since there would be on the order of $10^{15}$ entries.

The spatial dependence matrices, however, provide a way of escape. In using these matrices, it was observed that they are heavily weighted along the diagonal with decreasing entries farther from the diagonal. Figure 4 gives an
FIGURE 3-a.

\[
\begin{array}{ccc}
0 & 0 & 1 \\
0 & 1 & 1 \\
0 & 2 & 2 \\
2 & 3 & 3
\end{array}
\]

FIGURE 3-b. The general form of any grey tone spatial dependence matrix for an image with integer grey tone values 0 to 3. \( \#(i,j) \) stands for number of times grey tones \( i \) and \( j \) are neighbors.

\[
\begin{array}{cccc}
\text{Grey Tone} & 0 & 1 & 2 & 3 \\
0 & \#(0,0) & \#(0,1) & \#(0,2) & \#(0,3) \\
1 & \#(1,0) & \#(1,1) & \#(1,2) & \#(1,3) \\
2 & \#(2,0) & \#(2,1) & \#(2,2) & \#(2,3) \\
3 & \#(3,0) & \#(3,1) & \#(3,2) & \#(3,3)
\end{array}
\]

\[
\begin{array}{c}
0^\circ \\
2 \quad 4 \quad 0 \\
1 \quad 0 \quad 6 \\
0 \quad 0 \quad 1
\end{array}
\]

\[
\begin{array}{c}
90^\circ \\
6 \quad 0 \quad 2 \\
0 \quad 4 \quad 2 \\
2 \quad 2 \quad 2 \\
0 \quad 0 \quad 2
\end{array}
\]

\[
\begin{array}{c}
135^\circ \\
2 \quad 1 \quad 3 \\
1 \quad 2 \quad 1 \\
3 \quad 1 \quad 0 \\
0 \quad 0 \quad 2
\end{array}
\]

\[
\begin{array}{c}
45^\circ \\
4 \quad 1 \quad 0 \\
1 \quad 2 \quad 2 \\
0 \quad 2 \quad 4 \\
0 \quad 0 \quad 1
\end{array}
\]

FIGURE 3-c.

FIGURE 3-d.

FIGURE 3-e.

FIGURE 3-f.

1986
FIGURE 4. Example of Nearest Neighbor Grey Tone Dependence Matrices, Taken from Processing ERTS Data.
example of these matrices. Note that in each case the number of entries decreases as we move away from the diagonal. This suggests that neighboring resolution cells are similar. Choosing any resolution cell in an image at random, we are very likely to find nearly identical neighbors to the cell in all directions and less likely to find dissimilar neighbors. Clearly, a measure which indicates how similar the neighboring N-tuples are and how fast the similarity drops off with distance must contain textural information about the object imaged.

It is therefore reasonable to measure the difference between neighboring grey tone N-tuples and observe this distribution instead of computing the number of times each N-tuple neighbors every other N-tuple. In both cases we measure the co-occurrence of nearest neighbor grey tone N-tuples.

Since the textural features are based on the spatial dependence of grey tone N-tuples, our first step must be to define a binary relation between neighboring resolution cells on which the co-occurrence of grey tones N-tuples can be counted. As above, let \( L_x = \{1, 2, \ldots, N_x\} \) and \( L_y = \{1, 2, \ldots, N_y\} \) be the set of column and row indexes, respectively, so that \( ly \times lx \) is the set of resolution cells in the image. Let \( G = \{0, 1, \ldots, Ng\} \) be the set of possible grey tones that each component of every grey tone N-tuple can be assigned. Then, the image \( I \) can be defined by \( ly \times lx \times gy \times lg \times lx \).

Let \( R \) be the set of all pairs of resolution cells in a specified spatial relation. Then \( R \) is a binary relation on the set \( ly \times lx ; R \subseteq (ly \times lx) \times (ly \times lx) \). For example, the set of all distance 1 horizontally neighboring pairs of neighboring resolution cells would be defined by:

\[
R = \{(k, l), (m, n) \mid (ly \times lx) \times (ly \times lx) \mid k-m=0, |l-n|=1\}.
\]

The co-occurrence frequency of grey tone N-tuples \( (i_1, i_2, \ldots, i_N) \) and \( (j_1, j_2, \ldots, j_N) \) in spatial relation defined by \( R \) is

\[
P((i_1, \ldots, i_N), (j_1, \ldots, j_N)) = \frac{\# \{(k, l), (m, n) \mid R \mid I(k, l) = (i_1, \ldots, i_N), I(m, n) = (j_1, \ldots, j_N)\}}{\#R}
\]

where \( \# \) denotes the number of elements in the set.

Note that this \( R \) is symmetric. Assume that \( ((k, l), (m, n)) \) is in \( R \). Then \( k-m=0 \), and \( |l-n|=1 \) from the definitions of \( R \). But \( |l-n|=1 \) when \( |n-l|=1 \).
And if $|n-l|=1$ and $k-m=0$, then $(m, n), (k, l)$ is in $R$. Thus, $R$ is symmetric.

In fact, by the symmetry of any distance function, $R$, in general, must be symmetric. And since $R$ is symmetric, $P$ is also symmetric.

IV.3 Textural Feature Extraction Procedure

Let $R$ be a symmetric binary relation pairing nearby neighboring resolution cells. We define the frequency of grey tone N-tuple differences co-occurring in the spatial configuration defined by $R$ as

$$P(x_1, x_2, \ldots, x_N) = \frac{\# \{(i, j), (m, n) \in R \mid I(i, j) - I(m, n) = x_N\}}{\#R}.$$ 

Note that $P$ is an even function since

$$P(x_1, x_2, \ldots, x_N) = \# \{(i, j), (m, n) \in R \mid I(i, j) - I(m, n) = x_N\} / \#R$$

$$= \# \{(i, j), (m, n) \in R \mid I(i, j) - I(m, n) = -x_N\} / \#R$$

$$= \# \{(m, n), (i, j) \in R \mid I(m, n) - I(i, j) = -x_N\} / \#R$$

$$= P(-x_1, -x_2, \ldots, -x_N).$$

Referring to the monotonic behavior of nearly every column in the matrices of Figure 4, and assuming that this behavior occurs on every band of the ERTS multimag, it is reasonable to assume that the even frequency distribution $P(x_1, \ldots, x_N)$ of the nearby grey tone N-tuple differences can be adequately approximated using an ellipsoidally symmetric distribution; thus we may write

$$P(x_1, x_2, \ldots, x_N) = f(x' Ax)$$

for some monotonically decreasing function $f$. 

1939
FIGURE 5. Scattergram showing ellipsoidally symmetric distribution of differences for Bands 1 and 2 over a sample 64 x 64 ERTS image.

FIGURE 6. Histogram of the distribution of differences on Band 2 for distance 1 horizontally neighboring resolution cells.
This implies that only the function $f$ and the matrix $A$ need to be determined. We take $f$ to be one of the two forms $e^{-1/2 u^2}$, $(1+u^2)^{-m}$. Figure 5 is a scattergram of the differences for the first two bands of distance 1 horizontally neighboring resolution cells of a 64x64 sample image. Figures 5 and 6 clearly show the ellipsoidally symmetric functional form of the distribution of neighboring differences.

In Appendix I N-dimensional spherical coordinate systems and ellipsoidally symmetric distributions are discussed and it is shown that the matrix $A$ is proportional to the inverse covariance matrix of the N-tuple differences. Thus, we estimate $A$ by a matrix proportional to the inverse of the estimate for the covariance matrix.

Therefore, if the image is blocked into subimages of small area so that each subimage is essentially of one category, we can expect the distribution of grey tone N-tuple differences over each subimage to be a function only of the assumed form of the function $f$ and the covariance matrix of the difference vectors for grey tone N-tuples in a specified spatial relationship within the subimage. This leads us to consider textural features for multi-images based upon the elements of this spatial-spectral covariance matrix.

Consider each covariance matrix as a vector. Consider the distribution of the set of covariance matrices from the blocked image. Since the entries of the covariance matrix are the parameters of the distribution, we would like to have these entries invariant with respect to scale changes on the grey tone N-tuple differences. In order to do this, we scale the grey tone N-tuple differences so that all components have variance 1. The covariance matrix of these normalized differences is equivalent to the correlation matrix. Appendix II shows that this normalization procedure makes the covariance matrix invariant with respect to translating and scaling transformations on the grey tone N-tuples. The normalized covariance matrix can be considered as an extracted texture feature vector in an $N(N-1)/2$ dimensional hyperspace.
V. RESULTS OF CLASSIFICATION EXPERIMENTS ON ERTS MULTI-IMAGES

In order to obtain an initial estimate of performance of the multi-image texture features an experiment was performed on ERTS satellite imagery over Monterey Bay, California, image ID number 1002-18134. Using a small set of 64 sampled 32x32 subimages and training on 34 of these, 80 per cent of the remaining 30 test samples were correctly classified according to four land-use categories: coastal forest, annual grassland, urban area, and water, as shown in Table I. This is encouraging since previous accuracy using spatial dependence matrices on band 5 over the same general area was only 70.5 per cent as shown in Table II.

The ability to obtain good ground truth and several distinct categories in the California data was not the case for an ERTS image over Finney County, Kansas; which was used in the final experiments. Approximately a 40 mile by 60 mile section near Garden City, Kansas, on image (ID number) 1330-16512 was processed with initially four categories; grassland, large fields, small fields, and water. Both texture procedures, using spectral multi-image features and the single band 5 spatial dependence matrix, were used on the image. Tables III and IV show the results of classification for distance 1 resolution cells while Tables V and VI show distance 8 results. In both cases the single image classification is higher. However, when both distances 1 and 8 are used together, classification accuracy for both procedures is nearly identical, as shown in Tables VII and VIII, about 70 per cent.

This implies that much more information is contained in the Kansas data texturally than spectrally. In hope of adding more texture information, a measure of entropy relating to the correlation matrix was added to the spectral features. The effect of this new feature upon classification accuracy is presently being evaluated. We are also adding higher order components to the grey tone N-tuples in an attempt to gain more textural information from the data.
VI. CONCLUSION

The procedure developed here for the extraction of textural information from ERTS multi-images gives encouraging results. The procedure is quite simple to employ and is a natural evolution of the previous texture extraction technique based upon angular nearest neighbor grey tone dependence matrices. It retains the power of the previous approach to texture by characterizing the spatial inter-relationships, or co-occurrences, of the grey tone N-tuples present in a texture pattern in multi-images in such a way as to be invariant under monotonic grey tone transformations.
Kansas Data:

Image No.: 1330-16515 18 Jun 73
Area Processed: 40mi by 60mi (approx.)
Garden City, Kansas

Spectral-Textural Subimage:
32 points per line
32 lines

Textural Subimage:
64 points per line
64 lines
MSS Band 5

California Data:

Image No.: 1002-18134 25 Jul 72
Area Processed: 40 mi by 100 mi (approx.)
Monterey Bay, California

Spectral-Textural Subimage:
32 points per line
32 lines

Textural Subimage:
64 points per line
64 lines
MSS Band 5
### Table I. Contingency Table for Land Use Classification of ERTS-1 Satellite Imagery of Monterey Bay, California Using Spectral-Textural Features. Classification Accuracy on Test Set = 80%.

<table>
<thead>
<tr>
<th>TRUE CATEGORY</th>
<th>COASTAL FOREST</th>
<th>ANNUAL GRASSLAND</th>
<th>URBAN AREA</th>
<th>WATER</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>COASTAL FOREST</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>ANNUAL GRASSLANDS</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>URBAN AREA</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>WATER</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>TOTAL</td>
<td>5</td>
<td>8</td>
<td>12</td>
<td>5</td>
<td>30</td>
</tr>
</tbody>
</table>

### Table II. Results of Land Use Classification Experiments from Satellite Imagery of Monterey Bay, California.

<table>
<thead>
<tr>
<th>FEATURES</th>
<th>NO. OF SAMPLES IN TRAINING SET</th>
<th>NO. OF SAMPLES IN TEST SET</th>
<th>OVERALL ACCURACY OF TEST SET</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPECTRAL-TEXTURAL</td>
<td>34</td>
<td>30</td>
<td>80%</td>
</tr>
<tr>
<td>TEXTURAL</td>
<td>260</td>
<td>172</td>
<td>70.5%</td>
</tr>
</tbody>
</table>
### Table III.
Contingency Table for Land Use Classification of Satellite Imagery over Kansas Using Spectral-Textural Features, Distance 1, Four Categories. Classification Accuracy on Test Set = 67%.

<table>
<thead>
<tr>
<th>TRUE CATEGORY</th>
<th>GRASSLAND</th>
<th>LARGE FIELDS</th>
<th>SMALL FIELDS</th>
<th>WATER</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRASSLAND</td>
<td>47</td>
<td>11</td>
<td>5</td>
<td>0</td>
<td>63</td>
</tr>
<tr>
<td>LARGE FIELDS</td>
<td>3</td>
<td>142</td>
<td>32</td>
<td>1</td>
<td>178</td>
</tr>
<tr>
<td>SMALL FIELDS</td>
<td>17</td>
<td>46</td>
<td>52</td>
<td>0</td>
<td>115</td>
</tr>
<tr>
<td>WATER</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>TOTAL</td>
<td>67</td>
<td>200</td>
<td>95</td>
<td>2</td>
<td>364</td>
</tr>
</tbody>
</table>

### Table IV.
Results of Land Use Classification Experiments from Satellite Imagery of Kansas, Distance 1, Four Categories.

<table>
<thead>
<tr>
<th>FEATURES</th>
<th>NO. OF SAMPLES IN TRAINING SET</th>
<th>NO. OF SAMPLES IN TEST SET</th>
<th>OVERALL ACCURACY OF TEST SET</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPECTRAL-TEXTURAL</td>
<td>548</td>
<td>364</td>
<td>67%</td>
</tr>
<tr>
<td>TEXTURAL</td>
<td>140</td>
<td>88</td>
<td>76%</td>
</tr>
<tr>
<td>True Category</td>
<td>Assigned Category</td>
<td>Grassland</td>
<td>Large Fields</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------</td>
<td>-----------</td>
<td>--------------</td>
</tr>
<tr>
<td>Grassland</td>
<td>32</td>
<td>9</td>
<td>22</td>
</tr>
<tr>
<td>Large Fields</td>
<td>7</td>
<td>147</td>
<td>24</td>
</tr>
<tr>
<td>Small Fields</td>
<td>18</td>
<td>44</td>
<td>53</td>
</tr>
<tr>
<td>Water</td>
<td>0</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>57</td>
<td>204</td>
<td>103</td>
</tr>
</tbody>
</table>

**Table V.** Contingency Table for Land Use Classification of Satellite Imagery over Kansas Using Spectral-Textural Features, Distance 8, Four Categories. Classification Accuracy on Test Set = 64%.

<table>
<thead>
<tr>
<th>Features</th>
<th>No. of Samples in Training Set</th>
<th>No. of Samples in Test Set</th>
<th>Overall Accuracy of Test Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectral-Textural</td>
<td>548</td>
<td>364</td>
<td>64%</td>
</tr>
<tr>
<td>Textural</td>
<td>140</td>
<td>88</td>
<td>76%</td>
</tr>
</tbody>
</table>

**Table VI.** Results of Land Use Classification Experiments from Satellite Imagery of Kansas, Distance 8, Four Categories.
<table>
<thead>
<tr>
<th>TRUE CATEGORY</th>
<th>GRASSLAND</th>
<th>LARGE FIELDS</th>
<th>SMALL FIELDS</th>
<th>WATER</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRASSLAND</td>
<td>52</td>
<td>7</td>
<td>4</td>
<td>0</td>
<td>63</td>
</tr>
<tr>
<td>LARGE FIELDS</td>
<td>2</td>
<td>145</td>
<td>31</td>
<td>0</td>
<td>178</td>
</tr>
<tr>
<td>SMALL FIELDS</td>
<td>12</td>
<td>42</td>
<td>61</td>
<td>0</td>
<td>115</td>
</tr>
<tr>
<td>WATER</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>TOTAL</td>
<td>66</td>
<td>195</td>
<td>102</td>
<td>1</td>
<td>364</td>
</tr>
</tbody>
</table>

Table VII. Contingency Table for Land Use Classification of Satellite Imagery over Kansas Using Spectral-Textural Features, Distances 1 and 8, Four Categories. Classification Accuracy on Test Set= 71%.

<table>
<thead>
<tr>
<th>FEATURES</th>
<th>NO. OF SAMPLES IN TRAINING SET</th>
<th>NO. OF SAMPLES IN TEST SET</th>
<th>OVERALL ACCURACY OF TEST SET</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPECTRAL-TEXTURAL</td>
<td>548</td>
<td>364</td>
<td>71%</td>
</tr>
<tr>
<td>TEXTURAL (17)</td>
<td>140</td>
<td>88</td>
<td>73%</td>
</tr>
</tbody>
</table>

Table VIII. Results of Land Use Classification Experiments from Satellite Imagery of Kansas, using both Distances 1 and 8 of the Spectral-Textural Features and all 17 Textural Features.
### Table IX.
Contingency Table for Land Use Classification of Satellite Imagery over Kansas using Spectral-Textural Features with the BayesClassifier, Distance 7. Classification Accuracy of Test Set = 69%.

<table>
<thead>
<tr>
<th></th>
<th>Assigned Category</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GRASSLAND</td>
<td>LARGE FIELDS</td>
<td>SMALL FIELDS</td>
<td>WATER</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TRUE CATEGORY</td>
<td></td>
<td>49</td>
<td>4</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>LARGE FIELDS</td>
<td>150</td>
<td>27</td>
<td>1</td>
<td>184</td>
</tr>
<tr>
<td></td>
<td>SMALL FIELDS</td>
<td>34</td>
<td>67</td>
<td>2</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>WATER</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td>65</td>
<td>192</td>
<td>100</td>
<td>8</td>
</tr>
</tbody>
</table>

### Table X.
Contingency Table for Land Use Classification of Satellite Imagery over Kansas using Spectral-Textural Features with the BayesClassifier, Distance 8. Classification Accuracy of Test Set = 61%.

<table>
<thead>
<tr>
<th></th>
<th>Assigned Category</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GRASSLAND</td>
<td>LARGE FIELDS</td>
<td>SMALL FIELDS</td>
<td>WATER</td>
<td>TOTAL</td>
</tr>
<tr>
<td>TRUE CATEGORY</td>
<td></td>
<td>31</td>
<td>10</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>LARGE FIELDS</td>
<td>140</td>
<td>38</td>
<td>1</td>
<td>184</td>
</tr>
<tr>
<td></td>
<td>SMALL FIELDS</td>
<td>46</td>
<td>51</td>
<td>4</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>WATER</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td>49</td>
<td>230</td>
<td>110</td>
<td>6</td>
</tr>
</tbody>
</table>
APPENDIX I

N-DIMENSIONAL SPHERICAL COORDINATE SYSTEMS
AND ELLIPSOIDALLY SYMMETRIC DISTRIBUTIONS

We illustrate the N-dimensional spherical coordinate system in the calculation of the volume of the N-dimensional hypersphere. Next we show how suitable functions can be used to define ellipsoidally symmetric density functions and we determine the normalizing constant for any function. Finally, we show that for any ellipsoidally symmetric density \( f(\sqrt{x^T A x}) \), the matrix \( A \) is proportional to the inverse covariance matrix of \( x \) and we determine the constant of proportionality.

1.1 Volume of an N-dimensional Hypersphere

Let \( V \) be the volume of a N-dimensional hypersphere of radius \( r_o \). By definition

\[
v = \int \cdots \int dx_1 \, dx_2 \, \cdots \, dx_N \\
\sqrt{\sum_{l=1}^{N} x_l^2} \leq r_o
\]

To evaluate this N-fold integral, we transform to spherical coordinates.

\[
x_1 = r \cos \theta_1 \cos \theta_2 \cdots \cos \theta_{N-3} \cos \theta_{N-2} \cos \theta_{N-1} \\
x_2 = r \cos \theta_1 \cos \theta_2 \cdots \cos \theta_{N-3} \cos \theta_{N-2} \sin \theta_{N-1} \\
x_3 = r \cos \theta_1 \cos \theta_2 \cdots \cos \theta_{N-3} \sin \theta_{N-2} \\
\vdots \\
x_j = r \cos \theta_1 \cos \theta_2 \cdots \cos \theta_{N-3} \sin \theta_{N-j} \\
\vdots \\
x_N = r \sin \theta_1
\]

Figure 7 illustrates the geometry of the spherical coordinate system we use for a 3-dimensional system.
Three-Dimensional Spherical Coordinate System

\[ \begin{align*}
x_1 &= r \cos \theta_1 \cos \theta_2 \\
x_2 &= r \cos \theta_1 \sin \theta_2 \\
x_3 &= r \sin \theta_1
\end{align*} \]

Transformation between rectangular coordinate system and spherical coordinate system.

FIGURE 7. Three-Dimensional Spherical Coordinate System.
The Jacobian $J$ of this transformation is defined by the determinant $J$.

\[
J = \begin{vmatrix}
\frac{\partial x_1}{\partial r} & \frac{\partial x_2}{\partial r} & \cdots & \frac{\partial x_N}{\partial r} \\
\frac{\partial x_1}{\partial \theta_1} & \frac{\partial x_2}{\partial \theta_1} & \cdots & \frac{\partial x_N}{\partial \theta_1} \\
\vdots & \vdots & & \vdots \\
\frac{\partial x_1}{\partial \theta_{N-1}} & \frac{\partial x_2}{\partial \theta_{N-1}} & \cdots & \frac{\partial x_N}{\partial \theta_{N-1}}
\end{vmatrix}
\]

\[
\begin{vmatrix}
\cos \theta_1 \cos \theta_2 \cdots \cos \theta_{N-1} & \cos \theta_1 \cos \theta_2 \cdots \cos \theta_{N-2} \sin \theta_{N-1} & \cdots & \sin \theta_1 \\
-r \sin \theta_1 \cos \theta_2 \cdots \cos \theta_{N-1} & -r \sin \theta_1 \cos \theta_2 \cdots \cos \theta_{N-2} \sin \theta_{N-1} & \cdots & r \cos \theta_1 \\
-r \cos \theta_1 \sin \theta_1 \cdots \cos \theta_{N-1} & -r \cos \theta_1 \sin \theta_1 \cdots \cos \theta_{N-2} \sin \theta_{N-1} & \cdots & 0 \\
\vdots & \vdots & & \vdots \\
-r \cos \theta_1 \cos \theta_2 \cdots \sin \theta_{N-1} & r \cos \theta_1 \cos \theta_2 \cdots \cos \theta_{N-2} \cos \theta_{N-1} & \cdots & 0
\end{vmatrix}
\]

To find the value of the Jacobian, factor $r$ out of the last $(N-1)$ rows and from each column factor out its first entry.
$$J = r^{N-1} \cos \theta_1 \cos \theta_2 \cdots \cos \theta_{N-1} \sin \theta_1 \sin \theta_2 \cdots \sin \theta_{N-1}$$

$$\begin{array}{cccccc}
1 & 1 & 1 & \cdots & 0 \\
-tan \theta_1 & -tan \theta_1 & -tan \theta_1 & \cdots & cot \theta_1 \\
-tan \theta_2 & -tan \theta_2 & -tan \theta_2 & \cdots & 0 \\
\vdots \hspace{1cm} \vdots \hspace{1cm} \vdots \hspace{1cm} \cdots \hspace{1cm} \vdots \\
-tan \theta_{N-3} & \cdots & \cdots & \cdots & 0 \\
-tan \theta_{N-2} & cot \theta_{N-2} & \cdots & \cdots & 0 \\
-tan \theta_{N-1} & cot \theta_{N-1} & 0 & \cdots & 0
\end{array}$$

Subtracting column 2 from column 1, column 3 from column 2, \ldots, column N from column N-1 there results

$$\begin{array}{cccccc}
0 & 0 & 0 & \cdots & 0 & 1 \\
0 & 0 & 0 & \cdots & -tan \theta_1 & -cot \theta_1 & cot \theta_1 \\
\vdots \hspace{1cm} \vdots \hspace{1cm} \vdots \hspace{1cm} \cdots \hspace{1cm} \vdots \hspace{1cm} \vdots \hspace{1cm} \vdots \\
\vdots \hspace{1cm} \vdots \hspace{1cm} \vdots \hspace{1cm} \cdots \hspace{1cm} \cdots \hspace{1cm} \cdots \hspace{1cm} \cdots \\
-tan \theta_{N-3} & -cot \theta_{N-3} & \cdots & \cdots & \cdots & \cdots \\
-tan \theta_{N-2} & -cot \theta_{N-2} & cot \theta_{N-2} & \cdots & \cdots & \cdots \\
-tan \theta_{N-1} & -cot \theta_{N-1} & cot \theta_{N-1} & 0 & \cdots & 0 \\
\end{array}$$

Since all entries in the upper left triangle are zero, the value of the determinant is easily found as minus one times the product of entries on the lower left to upper right diagonal.
Notice that \( \tan \theta + \cot \theta = \frac{1}{\sin \theta \cos \theta} \). Now upon simplifying we obtain

\[
J = (-1)^N N^{N-1} \prod_{i=1}^{N-2} \cos \theta_i \cos \theta_1 \cos \theta_2 \cdots \cos \theta_{N-2}
\]

and \( |J| = r^{N-1} \cos \theta_1 \cos \theta_2 \cdots \cos \theta_{N-2} \) since

\[
\cos \theta_i > 0 \quad \text{for} \quad -\pi/2 \leq \theta_i \leq \pi/2, \quad i=1, 2, \ldots, N-2.
\]

In spherical coordinates the volume \( V \) of the \( N \)-dimensional hypersphere of radius \( r_0 \) is readily evaluated.

\[
V = \int_{r=0}^{r_0} \int_{\theta_1 = 0}^{\pi/2} \int_{\theta_2 = 0}^{\pi/2} \cdots \int_{\theta_{N-1} = 0}^{\pi/2} r^{N-1} \cos \theta_1 \cos \theta_2 \cdots \cos \theta_{N-2} \; dr \; d\theta_1 \cdots d\theta_{N-1}
\]

Separating the integrations,

\[
V = \int_{r=0}^{r_0} r^{N-1} dr \int_{\theta_1 = -\pi/2}^{\pi/2} \cos \theta_1 d\theta_1 \cdots \int_{\theta_{N-2} = -\pi/2}^{\pi/2} \cos \theta_{N-2} d\theta_{N-2} \int_{\theta_{N-1} = 0}^{\pi/2} d\theta_{N-1}
\]

Since

\[
\int_{-\pi/2}^{\pi/2} \cos \theta d\theta = \frac{\Gamma\left(\frac{N+1}{2}\right) \Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{N+2}{2}\right)}
\]

\[
\theta = \frac{\pi}{2}
\]

1954
\[ V = \frac{r_1^N}{N} \left[ \frac{\Gamma \left( \frac{N-1}{2} \right) \Gamma \left( \frac{1}{2} \right)}{\Gamma \left( \frac{N}{2} \right)} \right] \left[ \frac{\Gamma \left( \frac{N-2}{2} \right) \Gamma \left( \frac{1}{2} \right)}{\Gamma \left( \frac{N-1}{2} \right)} \right] \cdots \left[ \frac{\Gamma \left( \frac{2}{2} \right) \Gamma \left( \frac{1}{2} \right)}{\Gamma \left( \frac{3}{2} \right)} \right] 2^\pi \]

\[ = \frac{r_0^N}{N} \cdot \frac{2\pi}{\Gamma \left( \frac{N}{2} \right)} \Gamma \left( \frac{1}{2} \right) \Gamma \left( \frac{N-2}{2} \right) \frac{\Gamma \left( \frac{1}{2} \right)}{\Gamma \left( \frac{N-1}{2} \right)} \frac{\Gamma \left( \frac{2}{2} \right) \Gamma \left( \frac{1}{2} \right)}{\Gamma \left( \frac{3}{2} \right)} \]

But \( \Gamma \left( \frac{1}{2} \right) = \pi^{1/2} \) and \( \Gamma(1) = 1 \).

\[ V = \frac{r_0^N}{N} \cdot \frac{2\pi^{\frac{N-2}{2}}}{\Gamma \left( \frac{N}{2} \right)} = \frac{2r_0^N}{N} \cdot \frac{\pi^{\frac{N}{2}}}{\Gamma \left( \frac{N}{2} \right)} . \]
1.2 Suitable functions for Ellipsoidally Symmetric Distribution.

Suppose $f$ is a real function, defined on domain $\mathbb{R}$, a subset of $[0, \infty]$, and satisfying $f(u) \geq 0$ for all $u$ in $\mathbb{R}$ and $\int u^k f(u) du$ is finite for $k \leq N + 1$. We show that $f$ is suitable for defining an ellipsoidally symmetric density function and we determine the constant $c$ so that of $(\sqrt{x^t A x})$ is an ellipsoidally symmetric density.

Let $A$ be a $N \times N$ symmetric positive definite matrix and $X$ an $N \times 1$ vector. Consider the ellipsoidally symmetric function $f(\sqrt{x^t A x})$. We wish to determine a constant $C$ such that of $(\sqrt{x^t A x})$ is a density function.

It is clear that $C = \frac{1}{\int \cdots \int \frac{f(\sqrt{x^t A x})}{\sqrt{x^t A x}} \ dx_1 \cdots dx_N}$.

To determine the value of the integral, we will make a transformation which rotates and scales. Let $T$ be an orthonormal matrix such that $T^t A T = D$, where $D$ is a diagonal matrix. Make the change of variables

$$X = T D^{-1/2} z.$$ 

The Jacobian $J$ of this transformation is

$$J = \begin{vmatrix}
\frac{\partial x_1}{\partial z_1} & \frac{\partial x_2}{\partial z_1} & \cdots & \frac{\partial x_N}{\partial z_1} \\
\frac{\partial x_1}{\partial z_2} & \frac{\partial x_2}{\partial z_2} & \cdots & \frac{\partial x_N}{\partial z_2} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial x_1}{\partial z_N} & \frac{\partial x_2}{\partial z_N} & \cdots & \frac{\partial x_N}{\partial z_N}
\end{vmatrix} = |TD^{-1/2}| = |T| |D|^{-1/2}$$

1956
Since $T$ is an orthonormal matrix, $|T| = 1$ and
\[ |D| = |T^\top A T| = |T^\top| |A| |T| = |A|. \]
So the Jacobian is the determinant $|A|^{-1/2}$ which is positive since $A$ is positive definite.
\[
1 = \int \ldots \int f \left( \sqrt{x^\top A x} \right) \, dx_1 \ldots dx_N
\]
\[ \sqrt{x^\top A x} \in \mathbb{R} \]
\[
= |A|^{-1/2} \int \ldots \int f \left( \sqrt{z^\top D^{-1/2} T^\top A T D^{-1/2} z} \right) \, dz_1 \ldots dz_N
\]
\[ \sqrt{z^\top D^{-1/2} T^\top A T D^{-1/2} z} \in \mathbb{R} \]
\[
= |A|^{-1/2} \int \ldots \int f(z) \, dz_1 \ldots dz_N
\]
\[ \sqrt{z^\top z} \in \mathbb{R} \]

Now change to spherical coordinates.
\[
z_1 = r \cos^3 1 \cos^3 2 \ldots \cos^3_{N-1}
\]
\[
z_2 = r \cos^3 1 \cos^3 2 \ldots \cos^3_{N-2} \sin^3_{N-1}
\]
\[
\vdots
\]
\[
z_j = r \cos^3 1 \ldots \cos^3_{N-j} \sin^3_{N-j+1}
\]
\[
\vdots
\]
\[
z_N = r \sin^3 1
\]
The Jacobian of this transformation is \((-1)^{N-1} N^{-1} \cos^{N-2} \theta_1 \cos^{N-2} \theta_2 \ldots \cos^{N-2} \theta_{N-2}\).

\[
I = |A|^{-1/2} \int_{\theta_1=-\pi/2}^{\pi/2} \int_{\theta_2=-\pi/2}^{\pi/2} \ldots \int_{\theta_{N-2}=-\pi/2}^{\pi/2} \int_{\theta_{N-1}=0}^{\pi/2} f(r) \prod_{j=1}^{N-1} \cos^{N-2} \theta_j \cos^{N-3} \theta_2 \ldots \cos^{N-2} \theta_{N-1} \cos \theta d\theta_1 d\theta_2 \ldots d\theta_{N-1}
\]

\[
= |A|^{-1/2} \int r^{N-1} f(r) dr \int \cos^{N-2} \theta_1 d\theta_1 \int \cos^{N-3} \theta_2 d\theta_2 \ldots \int \cos^{N-2} \theta_{N-1} d\theta_{N-1}
\]

Since \(\int_{\theta=-\pi/2}^{\pi/2} \cos^{N-2} \theta d\theta = \frac{\Gamma \left(\frac{N+1}{2}\right)}{\Gamma \left(\frac{N+2}{2}\right)}\), the integrals are readily evaluated.

1958
\[ I = |A|^{-1/2} \int_{\mathbb{R}} r^{N-1} f(r) \, dr \left[ \frac{\Gamma\left(\frac{N-1}{2}\right)}{\Gamma\left(\frac{N}{2}\right)} \frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{N-1}{2}\right)} \frac{\Gamma\left(\frac{N-2}{2}\right)}{\Gamma\left(\frac{3}{2}\right)} \cdots \frac{\Gamma\left(\frac{2}{2}\right)}{\Gamma\left(\frac{1}{2}\right)} \right] 2\pi \]

\[
= |A|^{-1/2} \int_{\mathbb{R}} r^{N-1} f(r) \, dr \frac{\Gamma\left(\frac{N-2}{2}\right)}{\Gamma\left(\frac{N}{2}\right)} \frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{N}{2}\right)} 2\pi
\]

\[
= \frac{2(\pi)^{N/2}}{|A|^{1/2} \Gamma\left(\frac{N}{2}\right)} \int_{\mathbb{R}} r^{N-1} f(r) \, dr
\]

Therefore, the constant \( c \) is

\[
c = \frac{|A|^{1/2} \Gamma\left(\frac{N}{2}\right)}{2 \ (\pi)^{N/2} \int_{\mathbb{R}} r^{N-1} f(r) \, dr}.
\]
Next we determine the normalizing constant \( c \) for the forms \( e^{-\frac{1}{2}u^2} \) and \( (1 + u^2)^{-m} \).

**Case 1. Multivariate Normal**

The density function for the multivariate normal distribution is of the form

\[
f(v^\top A x) = e^{-\frac{1}{2}v^\top A x}, \quad 0 \leq v^\top A x \leq \infty
\]

or, \( f(r) = e^{-\frac{1}{2}r^2}, \quad 0 \leq r \leq \infty \).

Since

\[
\int_{r \in \mathbb{R}} f(r) dr = \int_{0}^{\infty} N^{-1} e^{-\frac{1}{2}r^2} dr = \int_{0}^{\infty} \left(2u\right)^{-\frac{N-2}{2}} e^{-u} du = \frac{N-2}{2} \Gamma \left(\frac{N}{2}\right)
\]

then the normalizing constant is

\[
c = \frac{\Gamma \left(\frac{N}{2}\right)}{2\pi^{\frac{N}{2}} |A|^{-\frac{1}{2}} \int_{r \in \mathbb{R}} N^{-1} f(r) dr} = \frac{\Gamma \left(\frac{N}{2}\right)}{2\pi^{\frac{N}{2}} |A|^{-\frac{1}{2}} \frac{N-1}{2} \Gamma \left(\frac{N}{2}\right)}
\]

and,

\[
f(\sqrt{v^\top A x}) = \frac{1}{(2\pi)^{\frac{N}{2}} |A|^{-\frac{1}{2}}} e^{-\frac{1}{2}v^\top A x}, \quad 0 \leq v^\top A x \leq \infty.
\]

**Case 2. Multivariate Pearson Type VII.**

Let \( f(\sqrt{v^\top A x}) = (1 + v^\top A x)^{-m}, \quad 0 \leq v^\top A x \leq \infty \),

then,

\[
\int_{r \in \mathbb{R}} f(r) dr = \int_{0}^{\infty} N^{-1}(1 + r^2)^{-m} dr = \int_{0}^{1} \left(1 - \frac{u}{u}\right)^{N-2} u^m \frac{du}{2u^2}
\]

\[
= \frac{1}{2} \frac{\Gamma \left(\frac{N}{2}\right) \Gamma \left(m - \frac{N}{2}\right)}{\Gamma \left(m\right)} \quad , \quad m > \frac{N}{2}.
\]

And the normalizing constant is

\[
c = \frac{\Gamma \left(\frac{N}{2}\right) \Gamma \left(m\right)}{2\pi^{\frac{N}{2}} |A|^\frac{1}{2}} \frac{\Gamma \left(\frac{N}{2}\right) \Gamma \left(m - \frac{N}{2}\right)}{\Gamma \left(m\right)} = \frac{\Gamma \left(m\right)}{\pi^{\frac{N}{2}} |A|^\frac{1}{2} \frac{N}{2} \Gamma \left(m - \frac{N}{2}\right)}
\]

and,

\[
f(\sqrt{v^\top A x}) = \frac{\Gamma \left(m\right) |A|^\frac{1}{2}}{\pi^{\frac{N}{2}}} (1 + v^\top A x)^{-m}, \quad m > \frac{N}{2}, \quad 0 \leq v^\top A x \leq \infty.
\]

**I.3 Covariance Matrix for Multivariate Distributions.**

Given the density function \( f(\sqrt{v^\top A x}) \) we want to find the covariance matrix \( \Sigma \),

\[
\Sigma = E(xx^\top) = c \int \ldots \int xx^\top f(\sqrt{v^\top A x}) dx_1 \ldots dx_N
\]

\( \sqrt{v^\top A x} \in \mathbb{R} \)

1960
where \( c \) is a normalizing constant and \( N \) is the dimension of \( x \). Using the orthonormal transformation \( T'AT = D \), where \( D \) is a diagonal matrix, and scaling with \( x = T D^{-1/2} z \), we have

\[
x = c \int \cdots \int (T D^{-1/2} z) (D^{-1/2} T') f(\sqrt{z'z}) |A|^{-1/2} dz_1 \cdots dz_N / \sqrt{z'z} \in \mathbb{R}
\]

since

\[
x'AX = z' D^{-1} T' A T D^{-1} z = z' D^{-1} D D^2 z = z'z
\]

and

\[
J = \begin{vmatrix}
\frac{\partial x_1}{\partial z_1} & \cdots & \frac{\partial x_N}{\partial z_1} \\
\vdots & \ddots & \vdots \\
\frac{\partial x_1}{\partial z_N} & \cdots & \frac{\partial x_N}{\partial z_N}
\end{vmatrix} = |T D^{-1/2} T'| = |D|^{-1/2} = |T' AT|^{-1/2} = |A|^{-1/2}.
\]

Rearranging,

\[
\frac{x}{\sqrt{x'x}} = c |A|^{-1/2} T D^{-1/2} \int \cdots \int z z' f(\sqrt{z'z}) dz_1 \cdots dz_N / \sqrt{z'z} \in \mathbb{R}
\]

where \( z'z \) is an \( N \times N \) matrix. Looking at the off diagonal terms, for \( i \neq j \),

\[
\int \cdots \int z_i z_j f(\sqrt{z'z}) dz_1 \cdots dz_N = 0 \quad \text{since we are integrating an odd function over even limits.}
\]

For terms of \( z_i \) along the diagonal, for \( i=j \),

\[
\int \cdots \int z_i^2 f(\sqrt{z'z}) dz_1 \cdots dz_N = \int \cdots \int z_1^2 f(\sqrt{z'z}) dz_1 \cdots dz_N
\]

and changing to spherical coordinates,

\[
= \int_0^{\pi/2} \cdots \int_0^{\pi/2} \int_0^{2\pi} r^2 \cos^2 \theta_1 \cdots \cos^2 \theta_{N-1} f(r) r^{N-1} \cos \theta_j \cdots \cos \theta_{N-2} dr d\theta_1 \cdots d\theta_N
\]

\[
r e \in \mathbb{R}, \quad \theta_1 = \pi/2, \quad \theta_{N-2} = -\pi/2, \quad \theta_{N-1} = 0.
\]

\[
= \int_0^{\pi/2} f(r) r^{N+1} \cos \theta_1 \cdots \cos \theta_{N-1} d\theta_1 \cdots d\theta_{N-2} 2 \int_0^{\pi/2} \cos \theta_{N-1} d\theta_{N-1}
\]

Since \( \int_{-\pi/2}^{\pi/2} \cos^k \theta d\theta = \frac{\pi}{2} \left( \frac{1}{2} \right)^{k/2} \frac{(k+1)}{k} \frac{(k+2)}{k+1} \),

\[
1961
\]
\[
\int \frac{1}{z^2} f\left(z, \frac{z}{r^2} \right) \, dz_1 \cdots dz_N = \int_r^{N+1} f(r) \, dr \,
\frac{\Gamma \left( \frac{N+1}{2} \right) \Gamma \left( \frac{1}{2} \right)}{\Gamma \left( \frac{k+2}{2} \right) \Gamma \left( \frac{N+1}{2} \right)} \cdots \frac{\Gamma \left( \frac{4}{2} \right) \Gamma \left( \frac{1}{2} \right)}{\Gamma \left( \frac{3}{2} \right) \Gamma \left( \frac{2}{2} \right)} \,
2 \frac{\Gamma \left( \frac{3}{2} \right) \Gamma \left( \frac{1}{2} \right)}{\Gamma \left( \frac{2}{2} \right)} \int_r^{N+1} f(r) \, dr.
\]

From 1.2,
\[
c = \frac{\Gamma \left( \frac{N}{2} \right) \left| A_{\gamma} \right|^2}{2 \pi^2 \int_r^{N-1} f(r) \, dr}
\]
so that
\[
\Psi = \frac{\Gamma \left( \frac{N}{2} \right) \left| A_{\gamma} \right|^2}{2 \pi^2 \int_r^{N-1} f(r) \, dr} \left| A \right|^{-\frac{1}{2}} TD^{-\frac{1}{2}} \frac{2\pi^2}{N \Gamma \left( \frac{N}{2} \right)} \int_r^{N+1} f(r) \, dr \, D^{-\frac{1}{2}} T
\]
and,
\[
\Phi = T \, D^{-1} \, T
\]
where \( T \, A \, T = D \), or \( D^{-1} = T \, A^{-1} \, T \).

So that
\[
\Psi = A^{-\frac{1}{2}} \frac{\int_r^{N+1} f(r) \, dr}{\left( \int_r^{N-1} f(r) \, dr \right)^{\frac{1}{2}}}
\]

Since the integrals are constants for any \( f \), the covariance matrix is directly proportional to \( A^{-1} \). We determine the constant of proportionality for the multivariate normal and Pearson Type VII distributions.
Case 1. Multivariate Normal

For the multivariate normal, the density is of the form \( f(r) = ce^{-\frac{1}{2}r^2} \) where
\( r^2 = x'Ax, \quad 0 \leq r < \infty \).

Since \( \int_{0}^{\infty} e^{-\frac{1}{2}r^2} \, dr = 2^{\frac{k-1}{2}} \Gamma \left( \frac{k+1}{2} \right) \),

then \( \int_{0}^{r} f(r) \, dr = c2^{\frac{N}{2}} \Gamma \left( \frac{N+2}{2} \right) \)

and \( \int_{r}^{\infty} f(r) \, dr = c2^{\frac{N-2}{2}} \Gamma \left( \frac{N}{2} \right) \).

so that \( \hat{x} = A^{-1} \frac{N}{2} \frac{N}{2} \Gamma \left( \frac{N}{2} \right) \)

and thus, \( \hat{x} = A^{-1} \).

or, \( A = \hat{x}^{-1} \).
Case 2  Multivariate Pearson Type VII.

\[ f(r) = c (1 + r^2)^{-m}, 0 \leq r \leq \infty. \]

Since

\[ \int r^{N+1} f(r) \, dr = \frac{c}{2} \frac{\Gamma\left(\frac{N}{2} + 1\right) \Gamma\left(m - \frac{N}{2} - 1\right)}{\Gamma(m)} \]

and

\[ \int r^{N-1} f(r) \, dr = \frac{c}{2} \frac{\Gamma\left(\frac{N}{2}\right) \Gamma\left(M - \frac{N}{2}\right)}{\Gamma(M)} \]

then,

\[ \varphi = A^{-1} \frac{c N}{2} \frac{r \left(\frac{N}{2}\right) \Gamma\left(\frac{N}{2} - 1\right)}{2N \Gamma\left(\frac{N}{2}\right)} \frac{2 \Gamma\left(m - \frac{N}{2} - 1\right)}{c I\left(\frac{N}{2}\right) \Gamma\left(m - \frac{N}{2} - 1\right)} \]

\[ = A^{-1} \frac{1}{2 \left(m - \frac{N}{2} - 1\right)} \]

Hence,

\[ A = \frac{1}{2 \left(m - \frac{N}{2} - 1\right)} \]

1964
APPENDIX II

NORMALIZATION PROCEDURE TO MAKE COVARIANCE MATRIX INVARIANT UNDER TRANSLATING AND SCALING TRANSFORMATIONS

Let \( \mathbf{\Sigma}_x \) be a covariance matrix for the difference vectors of grey tone N-tuples in a specified spatial relationship within a subimage. We transform the covariance matrix to obtain the normalized covariance matrix \( \mathbf{\Sigma}'_y \) using \( y = D\mathbf{x} \), where \( \mathbf{x} \) is the difference vector and \( D \) is diagonal. Thus, assuming zero mean,

\[
\mathbf{\Sigma}'_y = \mathbf{E}(yy') = \mathbf{E}(D\mathbf{x}x'D') = D \mathbf{E}(xx')D' = D \mathbf{\Sigma}_x D \quad \text{since } D' = D.
\]

For normalization, we have

\[
d_{ii} = \frac{1}{\sqrt{\sigma_{ii}^2}}
\]

where \( \sigma_{ii} \) is the \( ii \)th element of \( \mathbf{\Sigma}_x \) and is the variance \( \sigma_i^2 \) of the \( i \)th component of \( \mathbf{x} \).

Assume that all grey tone N-tuples have a scale factor \( a \) and an additive factor \( c \) so that for N-tuples \( ax_1 + c \) and \( ax_2 + c \), the difference becomes

\[
y = (ax_1 + c) - (ax_2 + c) = a(x_1 - x_2).
\]

Hence, translational effects due to bias terms are cancelled but scaling effects are marked by the diagonal transformation \( y = A\mathbf{x} \) so that the elements of the covariance matrix become

\[
\mathbf{\Sigma}' = \mathbf{E}(yy') = \mathbf{E}(Axx'A') \quad \mathbf{\Sigma}' = A\mathbf{\Sigma}_x A
\]

1965
where $A$ is a diagonal matrix. We must show that $\mathbf{t}_N$, the normalized covariance matrix of $\mathbf{t}_x$, is identical to $\mathbf{t}_y$. Normalizing $\mathbf{t}_x$ we have

$$
\mathbf{t}_N = D \mathbf{t} D
$$

$$
= D \left( A \mathbf{t}_x A \right) D
$$

where $D$ is again diagonal but in this case,

$$
d_{ii} = \frac{1}{\sqrt{\sigma_{ii} a_{ii}}}
$$

with $a_{ii}$ the $ii^{th}$ element of diagonal matrix $A$. For the $ij^{th}$ element of $\mathbf{t}_N$ we have

$$
\sigma_{Nij} = d_{ii} a_{ii} \sigma_{ij} a_{jj} d_{jj}
$$

$$
= \frac{a_{ii} \sigma_{ij} a_{jj}}{\sqrt{\sigma_{ii} a_{ii}} \sqrt{\sigma_{jj} a_{jj}}}
$$

$$
= \sigma_{ij}
$$

Thus, this procedure of normalization makes the entries of the normalized covariance matrix invariant with respect to translating and scaling transformations on the grey tone $N$-tuples.
REFERENCES


