Reprinted from

Ninth International Symposium

Machine Processing of

Remotely Sensed Data

with special emphasis on

Natural Resources Evaluation

June 21-23, 1983

Proceedings

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

GENERALIZED TEXTURE MEASURES FOR CLASSIFICATION AND IMAGE QUALITY ASSESSMENT OF REMOTE SENSING IMAGES

J.C.A. VAN DER LUBBE

National Aerospace Laboratory/Space Division/Remote Sensing Department
Amsterdam, The Netherlands

ABSTRACT

Statistical texture measures are frequently applied to texture analysis of remote sensing imagery. In this paper two generalized classes of statistical measures are introduced, which can be used as texture measures and which include entropy and angular second moment as special cases. These generalized texture measures have interesting properties, which are studied. Their relation with actual texture in the (sub)image is considered.

The generalized texture measures are applied for the classification of land use categories in remote sensing images. Special attention is paid to the classification by means of so-called texture images. It is shown that by suitable substitutions of the parameters into the generalized texture measures alternatives can be obtained for the entropy, which are more attractive with respect to implementation.

Finally, it is discussed how far statistical texture measures can play a role for the assessment of image quality of remote sensing imagery.

I. INTRODUCTION

It is well-known that texture analysis can be applied fruitfully to the segmentation and classification of remote sensing imagery for example with respect to land use or geological terrain types. Contrary to other image features like tone and colour, which are related to individual pixels, texture is a local phenomenon, which is concerned with more than one pixel in general.

There are several approaches for the extraction of textural features from remote sensing images. One of the common methods describes texture by means of the gray level co-occurrence matrix. This gray level co-occurrence matrix (GLCM) consists of the relative frequencies of pairs of pixels within a (sub)image or image block, which are characterized by their mutual distance and their gray level values. Generally distinction is made between horizontal, vertical and (cross)-diagonal oriented pairs of pixels. There is a close connection between the texture of the (sub)-image and the structure of the GLCM's. The various textural features of the (sub)image can be characterized by applying so-called texture measures, of which the values are related to the way in which the elements of GLCM's are spread out over these matrices. Besides the extraction of textural features from a GLCM, they can also be derived from gray level difference and sum histograms (GLDH and GLSH, respectively), which express the distributions of the gray level differences and sums of the pairs of pixels and which are directly related to GLCM. A survey of statistical texture measures for GLCM as well as GLDH and GLSH are given by Haralick et al.[5]

In this paper two generalized classes of texture measures will be introduced. These include the entropy and angular second moment, which are used as textural measures in literature, as special cases. The introduced measures are closely related to measures developed within information theory and that part of statistics which is concerned with the measurement of concentration and diversity. The properties of these generalized measures will be investigated. Special attention will be paid to the influence of the extra parameters on their properties and on their relation to the actual textural phenomena in the (sub)image.

In order to elucidate the usefulness of these measures for the classification of remote sensing imagery experiments were performed at NLR's RESEDA-system (a system for the digital processing of remote sensing data). This was done both for classification of image blocks, which are each assumed to belong to the same category, and for the classification of remote sensing imagery by means of so-called texture images.

Although texture measures are mostly applied for the classification and segmentation of images, it will be shown here that statistical texture measures (including the ones developed here) can also be used for the assessment of image quality and as such are a useful tool for the evaluation of imaging instruments and image processing techniques. This will be illustrated on the basis of an evaluation of the effect of a data compression technique, presently being studied by NLR under ESA-contract, on texture in remote sensing images.

II. CONCENTRATION MEASURE

The statistical texture measures in literature all characterize some properties concerning GLCM, GLDH and GLSH. The contrast and inverse difference moment measures express to which extent the elements of GLCM are scattered around the main-diagonal.[5] As such they measure the amount of local variations present in a (sub)image. Two other well-known texture measures are entropy (ENT) and angular second moment (ASM).[5] They characterize the (un)evenness of distributions and can be applied to the two-dimensional histogram of pairs of pixels as well as to the histograms of gray level difference and sum and are therefore measures for the (in)-homogeneity of the (sub)image.

The measure ENT originates from the statistical information theory and has been introduced by Shannon as early as 1948.[12] In the fifties ASM was already being applied within statistics as a measure of concentration in order to quantify the unevenness with which the features of elements of a set are distributed over these elements. Shannon's measure of entropy was used there as a measure of diversity or evenness. However, within these disciplines other measures, which can be considered as alternatives for ENT and ASM, have also been introduced over the years. Recently, a unifying approach has been given by the author with respect to both information theory and that part of statistics, which deals with the measurement of concentration and diversity.[13,14]
These generalized approaches are important not only for the mentioned scientific fields, but they can also give rise to generalized texture measures, which still include the measures ENT and ASM as special cases and which have interesting properties and advantages in comparison with the latter ones. Firstly, the generalized concentration measure will be introduced.

Assume n is the number of pixel pairs of an image block, of which the texture is analysed. These n pairs of pixels are divided into subsets on the basis of features which are informative with respect to texture, e.g. their individual gray level values and their gray level differences or sums. Let the number of subsets be equal to k. The set N is given by $N = (n_1,\ldots,n_k)$, whereby n_i, $i = 1,\ldots,k$ is the number of pairs of pixels belonging to subset x_i. Now, the concentration measure, which can express one aspect of the texture of the image block, should measure the unevenness with which the pairs of pixels are distributed over the k subsets. As a matter of fact, the most elementary measure of concentration is a function which is reversed proportional to the number of subsets k. For, if k is large, then the number of subsets is large and thus the concentration is small. However, in this approach the relative size of each subset is not taken into account. For that very reason it is preferable to base the concentration measure on the relative frequencies of the subsets. The ASM measure, given by

$$f_k(N) = \sum_{i=1}^{k} (n_i/n)^2,$$

is such a concentration measure. It can be interpreted as the probability that two randomly chosen pairs of pixels of the image block belong to the same subset. This measure can be generalized to

$$G_k(N) = G_k(n;n_1,\ldots,n_k) = \left\{\sum_{i=1}^{k} (n_i/n)^\rho\right\}^\sigma,$$

where $(\rho,\sigma) \in D$ and D is given by

$$D = \{(\rho,\sigma)\,|\,0 < \rho < 1, \sigma < 0 \cup \rho > 1, \sigma > 0\}.$$

With respect to the definition domain D it is mentioned that D is such that $G_k(N)$ satisfies concentration-like properties.

It is remarked that for $\rho = 2$ and $\sigma = 1$ ASM is obtained. Some other special cases of $G_k(N)$ which are important in the fields of statistics and information theory are now given below.

a. CONCENTRATION OF TYPE R, $(R > 1)$: $\sum_{i=1}^{k} (n_i/n)^R$,

b. R-MEAN CONCENTRATION, $(R > 0, R \neq 1)$:

$$\left\{\sum_{i=1}^{k} (n_i/n)^R\right\}^{1/(R-1)},$$

c. R-NORM CONCENTRATION, $(R > 1)$: $\left\{\sum_{i=1}^{k} (n_i/n)^R\right\}^{1/R}$.

These measures have interesting properties with respect to texture analysis as will be shown in the following sections.

III. PROPERTIES OF THE CONCENTRATION MEASURE

In considering literature on texture analysis it is noticed that less attention has been paid to the algebraic and analytic properties of the statistical texture measures (this also holds for ASM and ENT), whereas these properties can clarify which textural phenomena in the image block are described by the various texture measures. In general, the properties are only presented in terms of distributions, whereby it is insufficiently realized that the distributions concerning GLCM, GLDH or GLSH are consequences of the structure of the image block. It can be proved for example, that theoretically derived conditions for the extremes of the texture measures can never be satisfied in practice, due to the fact that for some given n and k no corresponding image block can be constructed. This should be noted by the interpretation of values of texture measures in an absolute sense. The properties of $G_k(N)$ have been extensively studied by the author.[13,14] However, only the most important ones are summarized in this paper. First we consider three algebraic properties of the concentration measure.

(i) $G_k(N)$ is invariant for a proportional change of the subsets.

(ii) $G_k(N)$ is a symmetric function of $n_1,\ldots,n_k$.

(iii) $G_k(N)$ is expansible.

It can be deduced from (i) that the concentration of an image block which is built up from mutually equal textural areas is equal to the concentration of one separate textural area. This implies that the concentration of an image block is determined by the texture of the smallest textural area, which still can be considered as an entity. Property (ii) is a characteristic property of all statistical texture measures. It holds that the concentration measure is independent of the sequence in which the pairs of pixels are considered. In fact this implies that other spatial information than the one concerning pairs of pixels is neglected. Property (iii) shows that empty subsets do not contribute to the value of the concentration measure.

Now, the maximum and minimum properties are presented.

(iv) $G_k(N) \leq G_k(n;n,0,\ldots,0) = 1.$

(v) $G_k(N) \leq G_k(n;n/k,\ldots,n/k) = k^{(1-\rho)\sigma}.$

Thus the maximum is achieved if all the pairs of pixels of an image block belong to the same subset. With respect to GLDH this implies that the image block contains pairs of pixels, which all have the same gray level difference.

It follows that the minimum is achieved if the pairs of pixels are equally distributed over the k subsets. This is only possible if n/k is an integer. Considering the minimum as a function of k, an image block with (k + 1) equal subsets will have a smaller concentration than an image block with k equal subsets. It may be concluded that the concentration measure not only depends on the unevenness of the distribution of the subsets but also on the number of subsets.
The smallest value is achieved for k = n and is equal to $n^{(1-\rho)\sigma}$. With respect to GLDM this means that all pairs of pixels of an image block belong to different subsets.

However, if e.g. the number of pairs of pixels is large in comparison with the number of subsets, then generally pairs of pixels cannot equally be divided over the subsets. In this case the minimum value which the concentration measure can achieve is larger than the one given above. This is expressed in the following inequality.

(vi) For $n = \alpha k + \beta$, where α and β are nonnegative integers, it yields

$$G_k(N) \geq G_k(n;\alpha,\ldots,\alpha,\overbrace{\alpha+1,\ldots,\alpha+1}^{\beta})$$
$$= \{(k-\beta)(\alpha/n)^{\rho} + \beta\{(\alpha+1)/n\}^{\rho}\}^{\sigma}.$$

Thus this subminimum is achieved if the pairs of pixels are distributed as equally as possible over the subsets.

Finally, two ordering properties of the concentration measure are given. These give a better insight into their relation with texture.

(vii) $G_k(N)$ is a monotone function, or for $m \leq n$:

$$G_2(m;m-1,1) \leq G_2(n;n-1,1).$$

(viii) In all cases of N, where for all i, $i = 1,\ldots,k$: $n_i = \sum_{j=1}^{1} n_{ij}$ it can be derived that

$$G_{k1}(n;n_{11},\ldots,n_{k1}) \leq G_k(n;n_1,\ldots,n_k).$$

The property of monotonicity states that the concentration of an image block, where one subset consists of one pixel pair whereas the other subset includes the remaining ones, is smaller than the concentration of an image block where the number of pairs of pixels of the second subset is smaller. This may be expected, since the unevenness between the two subsets becomes less.

Property (viii) implies for GLDH, that the concentration of gray level differences increases if the number of gray levels is reduced e.g. in order to add to statistical reliability. Moreover, it follows that the values of the concentration measure for GLDH or GLSH are always larger than the corresponding ones for GLCM.

A systematical study of the algebraic, analytic, and ordering properties can contribute to insight in the relation between the concentration measure and the textural phenomena in the image. For more details we refer to Van der Lubbe.[14]

IV. INFLUENCE OF THE PARAMETERS

As a matter of fact the influence of the parameters is important in the view of practical applications, where a choice of the values of the parameters ρ and σ, $(\rho,\sigma) \in D$, should be made. The values of ρ and σ have a direct impact on the behaviour of the concentration measure. For that very reason, the concentration measure is studied here as a function of its parameters and some considerations are given concerning the choices of ρ and σ.

The concentration measure $G_k(N)$ as function of ρ and σ is decreasing in ρ and σ for $\rho > 1$ and $\sigma > 0$, and increasing on the remaining part of the definition domain.

Considering the absolute minimum and maximum of the concentration measure, it can be concluded that, for given $(\rho,\sigma) \in D$, the range of values of $G_k(N)$ depends on ρ and σ.
The sensitivity for erroneously estimated distributions depends directly on this range and thus on ρ and σ. For $(\rho-1)\sigma \to 0$ the range tends to zero. The range is maximal for $(\rho-1)\sigma \to \infty$.

The values of the parameters also have a direct impact on the discrimination between different textures. For example, let us assume that there are two image blocks, one with pairs of pixels which are equally distributed over k subsets and the other with pairs of pixels which are equally distributed over (k+1) subsets. It is studied when there is maximal discrimination between these two image blocks on the basis of the concentration measure. It can be proved that the ASM measure is most sensitive if the image blocks contain 2 and 3 subsets, respectively. The same holds for the R-MEAN CONCENTRATION measure. With respect to the R-NORM CONCENTRATION measure for large values of R, the maximal discrimination is achieved for small k. For the small values of R the maximal discrimination is obtained for images blocks with a large number of subsets. Due to the fact that in practice k is larger than 2 or 3, the R-NORM CONCENTRATION with a small value of R is more appropriate for the discrimination between this type of textural image blocks than e.g. ASM.

Both the R-MEAN and R-NORM CONCENTRATION measures have interesting limiting properties. Let n_{max} be the number of pairs of pixels of the largest subset, then

$$\lim_{\sigma \downarrow 0, \rho\sigma \to 1} G_k(N) = n_{max}/n.$$

Thus the R-MEAN and R-NORM CONCENTRATION measures tend to n_{max}/n for $R \to \infty$. It also holds for $\rho \downarrow 0$ that $G_k(N)$ tends to n^{σ}. Together with the fact that both the R-MEAN and R-NORM CONCENTRATION measures are monotonic functions of R, this implies that for the large values of R the R-MEAN and R-NORM CONCENTRATION measures are mainly influenced by n_{max} and that for the small values of R the amount of concentration is rather determined by the relative frequencies of the number of all the subsets. The R-MEAN CONCENTRATION measure tends to n_{max}/n as its lower bound, whereas the R-NORM CONCENTRATION measure tends to n_{max}/n as its upper bound. Due to this properties and the fact that they include n_{max}/n as a special case, (which is also used as a texture measure in literature), both the R-MEAN and R-NORM CONCENTRATION measures are useful texture measures; by suitable substitutions of the parameters different textural properties can be emphasized.

The behaviour of the R-MEAN and the R-NORM CONCENTRATION measures is illustrated in Figures 1 and 2. It also follows from this figures that the R-MEAN CONCENTRATION measure is rather insensitive to changes in even distributions for small values of R. The same holds for the R-NORM CONCENTRATION measure.

The final choice of the parameter values is determined by factors as the textures to be discriminated or the textural features to be extracted.

V. DIVERSITY MEASURES

The generalized concentration measure gave rise to the ASM measure. It will be shown here that ENT is closely related to generalized diversity measures. Diversity measures can be defined analogously to measures of concentration. It is a matter of fact that if the diversity of an image block is large then the concentration of this image block is small. Conversely, a set of pairs of pixels with a large concentration would have a small diversity. For that very reason, a diversity measure should be monotonically decreasing function of the concentration measure. Furthermore, because $G_k(N)$ is maximal (i.e. equal to 1) if the image block consists of just one type of pixel pair, it is reasonable to require for measures of diversity that in this case they become equal to zero. These starting points lead to the following definitions of measures of diversity.

Let $(\rho,\sigma,\delta) \in C$, where $C = \{\rho,\sigma,\delta \mid (\rho,\sigma) \in D \cup \delta > 0\}$ then the generalized diversity measures are given by:[13,14]

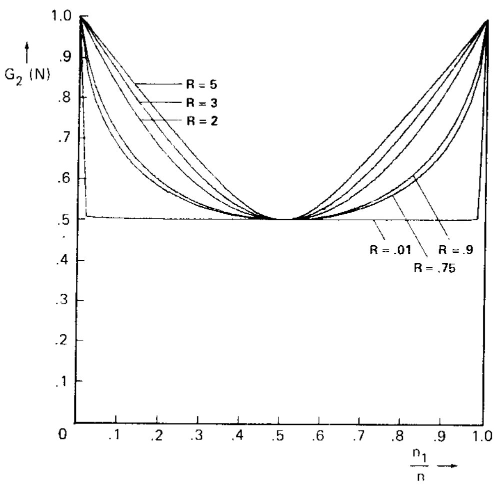

Figure 1. The R-MEAN CONCENTRATION as function of R.

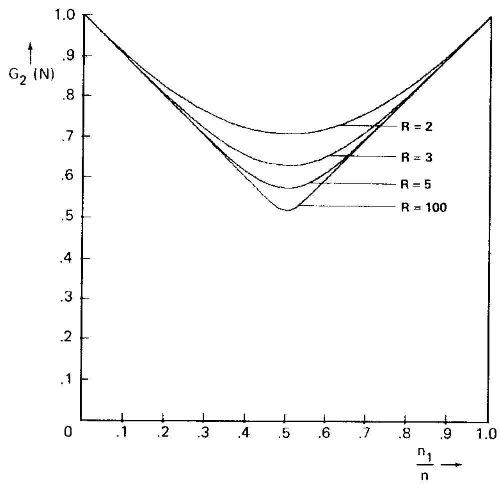

Figure 2. The R-NORM CONCENTRATION as function of R.

1. Logarithmic diversity: ${}^1H_k(N) = -\delta.\log[G_k(N)]$,

2. Linear diversity : ${}^2H_k(N) = \delta[1-G_k(N)]$,

3. Hyperbolic diversity : ${}^3H_k(N) = \delta[1/G_k(N)-1]$.

It can be easily seen that they are non-negative and that they are equal to zero in the case, where the image block represents just one type of pixel pairs. Furthermore, it follows that by using the minimum property of $G_k(N)$ the diversity measures achieve their maximum if the elements of an image block are distributed over the classes as equally as possible. Due to the relation between diversity and concentration the properties of the diversity measures can be directly derived from the properties of the measures of concentration. With respect to the ordering properties diversity and concentration have a reversed behaviour.

It can be concluded from an axiomatic characterization theorem, that these diversity measures are less arbitrary than they seem to be.[13] The introduced measures include the well-known information and diversity measures, and also ENT. A number of diversities will be given here. They can be obtained by using the definition of $G_k(N)$ and by suitable substitutions of ρ and σ.

a. The DIVERSITY OF ORDER R, $(R > 0,\ R \neq 1)$:[11]

$$H_k^a(N) = \frac{1}{1-R}\log\left[\sum_{i=1}^{k}(n_i/n)^R\right].$$

b. The R-NORM DIVERSITY, $(R > 0,\ R \neq 1)$:[1,2,13]

$$H_k^b(N) = R\left\{1-\left[\sum_{i=1}^{k}(n_i/n)^R\right]^{1/R}\right\}/(R-1).$$

It follows by taking the limit for $R\to 1$ that both $H_k^a(N)$ and $H_k^b(N)$ include Shannon's measure ENT given by

$$H_k^s(N) = -\sum_{i=1}^{k}(n_i/n)\log(n_i/n),$$

as a special case. It is an advantage of both the DIVERSITY OF ORDER R and the R-NORM DIVERSITY that they are decreasing in R. For increasing value of R the influence of the largest subset increases, until for $R \to \infty$ the measures only depend on n_{max}/n. This implies that by using these measures as texture measures the choice of the parameter R determines more or less the textural property, which is measured.

Both the DIVERSITY OF ORDER R and the R-NORM DIVERSITY measures are closely related to ENT, not only as limiting case but also as expressed in the following inequality which holds for $R > 1$

$$H_k^s(N) \geq H_k^a(N) \geq H_k^b(N).$$

For $0 < R < 1$ the inequality signs are reversed. The close relation between the DIVERSITY OF ORDER R, the R-NORM DIVERSITY and ENT suggests that, for small values of R they can be used as an alternative for ENT. The advantage is that this saves processing time.

One way to investigate this is tracing how far $H_k^a(N)$ and $H_k^b(N)$ are order preserving with respect to ENT. This means if $H_k^a(N_1)$ and $H_k^a(N_2)$ are the amounts of diversity for two image blocks and if $H_k^a(N_1) \geq H_k^a(N_2)$, then it should hold that $H_k^s(N_1) \geq H_k^s(N_2)$. It can now be proved that this is the case for $H_k^a(N)$ as well as for $H_k^b(N)$ and for all $R \in (R_1,R_2)$ where $R_1 \in [0,1)$ and $R_2 \in (1,\infty]$, and whereby R_1 and R_2 depend on the distributions of N_1 and N_2. Tests on remote sensing images have shown that in 98 % of all cases $H_k^a(N)$ and $H_k^b(N)$ are order preserving with respect to $H_k^s(N)$, which makes them attractive alternatives for the ENT measure.

Results with respect to classification are studied in section VII, where they are used in combination with other textural measures including the concentration measures.

VI. CONCENTRATION AND DIVERSITY CHANGE

In the foregoing sections measures of concentration and diversity have been introduced. Thereby the distribution of pairs of pixels for each image block was assumed to be fixed. However, when considering corresponding image blocks of multispectral imagery this in general, is not the case. For that very reason, it is of interest to define measures for concentration and diversity change. This is also important for those applications, where statistical measures are used for relative image quality assessment of remote sensing imagery, e.g. when one is interested in loss of texture due to subjection of remotely sensed data to data compression.

Analogously to Bruckmann an easy measure for change in concentration can be given by

$$E = G_k(N_2)/G_k(N_1),$$

where $G_k(N_1)$ is related to the reference image (the original one by data compression) and $G_k(N_2)$ is related to the image of which the texture loss is analyzed.[3] With the change of minimal concentration to maximal concentration, E becomes equal to $1/G_k(N_1)$. E becomes equal to $k^{(1-\rho)\sigma}/G_k(N_1)$ if the concentration decreases to a minimum.

Thus E is bounded by

$$k^{(1-\rho)\sigma} \leqq E \leqq k^{(\rho-1)\sigma}.$$

If one requires normalization of E within bounds of between -1 and +1, then one can choose

$$\mathrm{E} = (E-1)/E+1).$$

It easily follows that E is increasing in E. Moreover it holds that

$$-a \leqq \mathrm{E} \leqq a,$$

where $a = (k^{(\rho-1)\sigma}-1)/(k^{(\rho-1)\sigma}+1)$.

If $G_k(N_1) = G_k(N_2)$, E is equal to 1 and thus E takes on the value 0. The upper bound is achieved if the concentration changes from a minimum to a maximum. For instance, with respect to GLCM this would imply that an image block, which consists of all different pairs of pixels, would change into an image block, where all pixels have the same gray level value. The converse holds for the lower bound.

It is also possible to define measures for diversity change. However, these measures will not be considered here. For more details refer to literature.

It will be shown in section VIII that the measure of concentration change is suitable for image quality assessment.

VII. STATISTICAL TEXTURE MEASURES BY CLASSIFICATION

With respect to the classification of remotely sensed images on the basis of texture analysis two approaches can be followed.

First classification can occur with image blocks, which are considered a priori as uniform regions and in which the pixels are assumed to belong to the same category. Classification is now based on the extraction of both spectral and textural features from the separate image blocks. This method has proved to be successful for land use classification of image blocks derived from Landsat-MSS images.[6]

Another approach, which takes more account of the information content of the individual pixels, is based on textural transformations. Around each pixel of the image a pixel window of fixed dimension is assumed. Within a window textural measurements are performed and the corresponding values are assigned to the centre pixel of the window. By appropriate scaling so-called texture images can be obtained for each textural feature, which besides the spectral channels can be used as input for the classifier. Good results were obtained by Hsu for monospectral aerial images where the window consisted of (3 x 3) pixels.[8] Irons and Petersen have performed experiments concerning Landsat-MSS images, where windows were used with sizes of (3 x 3) and (5 x 5) pixels.[9] In both cases texture measures were applied that differed from the measures of Haralick or the here developed generalized ones.

Experiments concerning both approaches were carried out using NLR's RESEDA-system (for the digital processing of remote sensing data) in order to evaluate the classification results, by using statistical texture measures, including the here developed ones. With respect to the concentration and diversity measures various substitutions for the parameters were considered. With respect to classification based on textural transforms texture images were derived from Landsat-MSS images, band 7. Since it is well-known that textural features related to GLDH perform as well as the ones based on GLCM, experiments were performed only with the help of GLDH. The original image was firstly histogram equalized, whereby in order to guarantee sufficient statistical reliability the number of gray levels was reduced to 32. As a matter of fact the size of the pixel window has also a direct impact on the final classification results.

Large pixel windows may increase the statistical reliability of the computed textural features. However, a disadvantage of large pixel windows is that especially in the case of relatively small textural areas information of neighbouring regions is also included.

For our main experiments a window of (5 x 5) pixels was chosen. Furthermore, only pairs of pixels consisting of adjacent pixels were considered.

For each textural feature the results for a window were averaged with respect to the four directions, and this averaged value was assigned to the centre pixel. By appropriate scaling texture images were obtained.

From the texture images it could be derived that other textural features could be extracted by suitable substitutions of the parameters.

With respect to the R-NORM CONCENTRATION and DIVERSITY, the R-MEAN CONCENTRATION and the DIVERSITY OF ORDER R, the influence of the most frequent gray level difference within the window was clear for the large values of R; for the small values of R all gray level differences are taken equally into account. These measures were all discriminative for small as well as large values of R.

Evidently there is correlation between the various concentration and diversity measures. For example, the R-NORM CONCENTRATION and DIVERSITY measures are equivalent, if used for a maximum likelihood classification. This is due to the fact that the R-NORM DIVERSITY measure can be considered as a linear transformation of the R-NORM CONCENTRATION measure. Strong correlation also holds for the R-NORM and R-MEAN CONCENTRATION measures with large values of R, which is due to their limiting properties. However, application of concentration and diversity measures, used for both small and large values of the parameters, is useful, since different aspects of the textural phenomena within the window are enlightened. The CONCENTRATION OF TYPE R proved to be less successful for large values of R.

For small values of R the texture images based on the R-NORM DIVERSITY and the DIVERSITY OF ORDER R were highly correlated with Shannon's entropy measure. Due to the fact that textural transforms themselves take a long processing time and ENT is hardly computed because of the logarithmic function, the R-NORM DIVERSITY and the DIVERSITY OF ORDER R are attractive alternatives for ENT.

Classification was made using besides the separate channels a combination of textural images based on concentration or diversity measures (with small and large parameter value) and measures such as "inverse difference moment" and "difference mean"; since the latter ones were more related to the size of the differences than to the distribution of the differences.

The relative enhancement of classification results was most obvious when the texture images were used in combination with a single channel than with several spectral channels. Furthermore, supervised classification in regions which have distinct texture led to consistent labeling as compared with supervised classification based only on spectral information. Where pixels of training areas with distinct texture were not classified as members of the same category when only spectral features were used, the results were considerable better when textural features were extracted. It also seemed possible to discriminate between regions which had the same spectral signature.

However, some critical marginal notes should be made with respect to the application of textural transformations for the classification of images like Landsat-MSS images. For pixel windows of (3 x 3) or (5 x 5) pixels the statistical texture measures, including the ones introduced by Haralick et al. act more as edge detectors than as measures of texture. This holds particularly for a measure as the difference mean, which measures the averaged gray level difference over a window. This is especially inconvenient in image parts which are characterized by small bounded areas with different texture (compare the left side of the Landsat-MSS image in Figure 3). The edge effect appears in the classified image as rejected or misclassified pixels between adjacent classes. This can only be avoided at the cost of discrimination. This coincides with the results of Irons ans Petersen and seems inherent in the method of textural transforms.[9]

The problem of the edge effect did not occur in Landsat-MSS imagery with relatively large texture regions. Classification results based on spectral and texture images were considerably better for aerial photography, as shown by the experiments of Hsu, and for images e.g. from Thematic Mapper with higher spatial resolution.[8]

Finally, a few remarks are made with respect to the classification of image blocks according to the first approach. Experiments were carried out for the classification of image blocks of Landsat-MSS imagery, representing different land use categories, on the basis of spectral and textural information (GLDH). As might be expected the generalized concentration and diversity measures could be applied succesfully. By means of the choice of the parameter values the textural aspects to be emphasized by the concentration and diversity measures could be slightly controled. It was also established that the ENT measure could be replaced by the DIVERSITY OF ORDER R and R-NORM DIVERSITY measures, provided that the parameter R is small. Since concentration and diversity measures applied on GLDH, are more concerned with the distribution of the gray level differences than with the differences itselves, the best results were obtained if they were combined with texture measures as "inverse difference moment" and "difference mean".

With respect to both approaches generalized concentration and diversity measures proved to be successful. The final selection of the type of concentration and diversity measures to be applied to the classification of remote sensing images, as well as the choice of their parameter values depends on various factors like desired sensitivity for subtle changes in even or uneven distributions, the type of textures to be discriminated, the textural features to be observed and the computa-

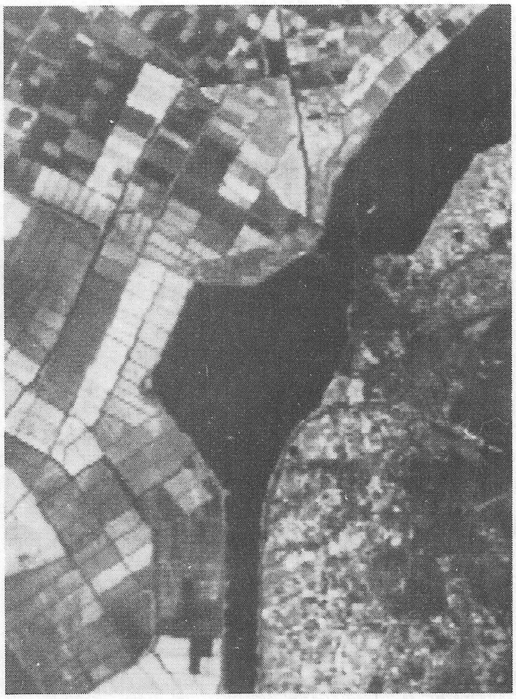

Figure 3. Landsat-MSS image (band 7) of the Harderwijk area, The Netherlands.

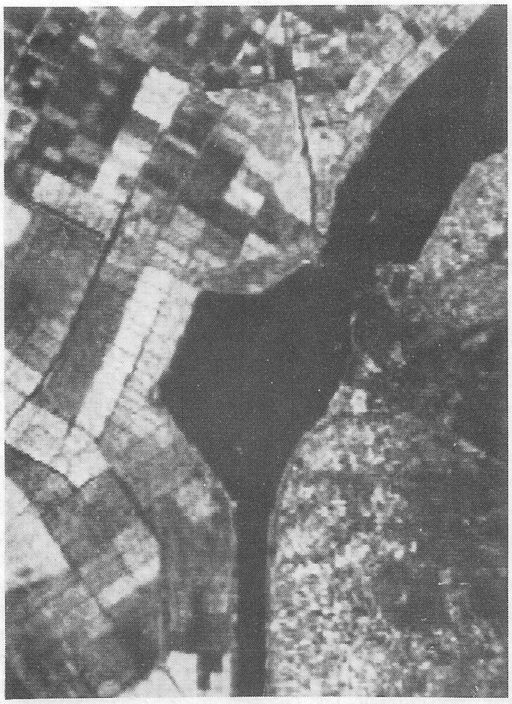

Figure 4. Landsat-MSS image (band 7) of the Harderwijk area, The Netherlands. Compression ratio 9.0.

bility. Hereby study of their properties as in the foregoing sections can assist.

VIII. STATISTICAL TEXTURE MEASURES AS IMAGE QUALITY CRITERIA

Due to recent developments and possibilities in the field of image generation and processing, the interest for and the importance of the assessment of image quality increases. As a matter of fact image quality can be determined with the help of subjects, who pass (subjective) judgement on the quality of images. However, since conditions such as experience of the interpreter, type of image, experimental conditions etc. play an important role with respect to the subjective image quality judgement, it is in general difficult to make conclusions concerning the effectiveness of systems and methods generating the judged images.

Quantitative measures like signal-to-noise ratio, mean squared error, are more objective and reproducable and are commonly used for assessment of image quality, but their drawback is that they coincide badly with what the human visual system considers as image quality. Furthermore, they give a rather global impression of image quality.

In general the interpreter and user of remote sensing imagery is interested in more detailed information with respect to the quality of remote sensing images. For a remote sensing user the quality of an image is also determined by aspects like classification accuracy, edge-sharpness (e.g. in view of cartographic applications) or texture-coarseness (e.g. in view of agricultural or silvicultural applications). For that very reason, it is useful to develop criteria for image quality, which are both quantitative and each emphasize a distinct aspect of image quality. Pratt [10] has already suggested that texture measures can perhaps be used for the assessment of that aspect of image quality which can be described as texture. Here, the statistical texture measures, including the ones developed here and the measures for concentration change are used for this purpose.

Table 1. Concentration changes by data compression. (compression ratios 5.5, 9.0; the measures are based on changes of the R-MEAN CONCENTRATION)

	water		grass-lands		agriculture	
	5.5	9.0	5.5	9.0	5.5	9.0
GLCM	0.729	0.790	-0.159	-0.163	-0.439	-0.443
GLCM	0.638	0.698	-0.287	-0.305	-0.500	-0.500
GLDH	0.408	0.454	-0.213	-0.200	-0.227	-0.221
GLDH	0.320	0.358	-0.315	-0.307	-0.235	-0.212
GLSH	0.477	0.578	-0.124	-0.092	-0.206	-0.260
GLSH	0.387	0.474	-0.143	-0.143	-0.208	-0.261

Table 2. Textural features and data compression. (compression ratios 0, 5.5, 9.0; COR = correlation between neighbouring pixels; IDM = inverse difference moment; ENT = Shannon's entropy)

	water		
	0	5.5	9
COR	0.11	0.38	0.49
IDM	0.456	0.874	0.893
ENT (GLCM)	4.202	2.131	1.460
ENT (GLDH)	2.114	1.026	0.843
ENT (GLSH)	2.820	1.830	1.294
	grass-lands		
	0	5.5	9
COR	0.52	0.52	0.53
IDM	0.103	0.073	0.078
ENT (GLCM)	8.074	8.409	8.418
ENT (GLDH)	4.812	5.126	5.086
ENT (GLSH)	6.358	6.606	6.581
	agriculture		
	0	5.5	9
COR	0.70	0.72	0.74
IDM	0.220	0.151	0.154
ENT (GLCM)	7.158	8.146	8.162
ENT (GLDH)	3.657	4.003	3.987
ENT (GLSH)	5.612	6.067	6.151

Dinstinction can be made between absolute and relative image quality assessment. In the first case the assessment concerns a single image, whereas in the second case a comparison is made between a pair of images. It is evident, that especially by the application of statistical texture measures for absolute image quality assessment the properties of the various measures should fully be taken into account in order to well interpret the numerical values of the texture measures.

With respect to textural transformation it was necessary, depending on the size of the window, to reduce the number of gray levels of the image before calculating the textural properties. From the view-point of image quality assessment this is less preferable. For instance, in comparing an image, which is distorted by data compression, with the original image, gray level reduction would lead to a considerable loss of information with respect to the factual influence of data compression on texture. Therefore, it is necessary to use a relatively large window for the computation of textural phenomena, in order to achieve sufficient statistical reliability. However, under influence of edges and boundaries in the image, too large window sizes can disturb the results. This holds especially true for a texture measure like "contrast". The "inverse difference moment" can be used as an alternative for this measure, since this measure expresses local variations in pixel gray levels, whereby a relatively large difference in gray levels contributes least to the total amount.

In the following example various texture measures are used for the evaluation of the impact of data compression on the image quality qua texture. Data compression was carried out on the Landsat-MSS image of figure 3. The applied data compression algorithm was based on one of the methods studied by NLR under an ESA-contract concerning the design and realization of an imaging sensor signal processing system for use on board scientific satellites. Following this method image blocks are transformed with the help of the discrete cosine transformation. The coefficients of the transformed image blocks are represented in a number of bits, according to a so-called bit map. In this example the RMS error per image block was fixed. The overall compression ratios were 5.5 and 9.0, whereby the RMS errors were equal to 5.2 and 6.6, respectively.

The reconstructed image for compression ratio 9.0 is given in figure 4.
Textural features now were computed within pixel windows of (20 x 20) pixels for both the original image and the (de)compressed image. In table 1 the concentration changes are given for three land-use categories: water, field agriculture, and grass-lands.

It can be concluded from table 1 that the concentration with respect to the category water increases for increasing compression ratio, the converse holds for the land use categories field agriculture and grass-lands. There concentration is lower for the (de)compressed images in comparison with the original image. In table 2 three other textural features are given.
The inverse difference moment feature expresses that for the categories field agriculture and grass-lands the contrast first increases for increasing compression ratio. After compression ratio 5.5 it decreases. Also now the behaviour of category water is reversed. From the entropy it can be concluded that with the exception of water the entropies of the pixel pairs, gray level differences and sums are larger in the (de)compressed image than in the original image.

Evidently, additional texture is generated in the (de)compressed image, with the exception of the water area. Texture measures can clarify the effect of the data compression method on the various aspects of texture. Also curves can be obtained analogously to rate distortion curves with respect to the mean squared error. The relation between these curves and the interpretability of the image is studied.

IX. CONCLUSIONS

In this paper generalized concentration and diversity measures were introduced which can be applied to texture analysis of remotely sensed imagery. The relation between their properties and the textural phenomena in the image itself was investigated. At the same time, it clarified the properties of known texture measures like angular second moment and entropy. It has been shown that the properties of the generalized concentration and diversity measures can be changed by appropriate substitutions of their parameters, whereby other textural characteristics can be extracted.
Moreover, some special cases of these generalized measures proved to be suitable alternatives for Shannon's entropy measure, both in the sense of computability, and of algebraic and analytic properties.

The here developed texture measures were useful for the classification of segments of remote sensing images.

Experiments were also performed concerning classification of Landsat-MSS images with the help of textural transformations based on statistical texture measures. It has been shown that this technique is useful for the classification of Landsat-MSS images with the exception of images which contain relatively small and bounded textural regions. Good results with textural transformations were obtained for imagery with high spatial resolution, e.g. aerial photography and Thematic Mapper imagery.

Finally, another application of statistical textures measures, including the ones here introduced was discussed. It has been shown that they can play a role in the assessment of image quality of remote sensing images and by the evaluation of imaging systems and image processing techniques, e.g. data compression techniques. This application is currently being further studied under contract for the Netherlands Agency for Aerospace Programs (NIVR).

X. REFERENCES

1. Arimoto, S. (1971), Information theoretical considerations on estimation problems, Inform. Contr., Vol. 19, pp. 181-194.

2. Boekee, D.E., Van der Lubbe, J.C.A. (1980), The R-norm information measure, Inform.Contr., Vol. 45, pp. 136-155.

3. Bruckmann, G. (1969), Einige Bemerkungen zur statistischen Messung der Konzentration, Metrika, Vol. 14, pp. 183-213.

4. Daroczy, Z. (1963), Über die gemeinsame Charakterisierung der zu den nicht vollständigen Verteilungen gehörigen Entropien von Shannon und von Renyi. Z. Wahrscheinlichkeitstheorie und verw. Gebiete, Band 1, pp. 381-388.

5. Haralick, R.M., Shanmugam, K., Dinstein, I. (1973), Textural features for image classification, IEEE Trans. Syst., Man, Cybern., Vol. SMC-3, pp. 610-621.

6. Haralick, R.M., Shanmugam, K.S. (1974), Combined spectral and spatial processing of ERTS imagery data, Remote Sensing of Environment, 3, pp. 3-13.

7. Haralick, R.M. (1979), Statistical and structural approaches to texture, Proc. IEEE, Vol. 67, No. 5, pp. 786-804.

8. Hsu, S. (1978), Texture-tone analysis for automated land-use mapping, Photogrammetric Engineering and Remote Sensing, Vol. 44, No. 11, pp. 1393-1404.

9. Irons, J.R., Petersen, G.W. (1981), Remote Sensing of Environment, 11, pp. 359-370.

10. Pratt, W.K. (1979), Image transmission techniques, New York.

11. Renyi, A. (1960), On measures of entropy and information. Proc. Fourth Berkeley Symp. Math. Statist. and Prob., No. 1, pp. 547-561.

12. Shannon, C.E. (1948), The mathematical theory of communication. Bell. Syst. Tech. J., Vol. 27, pp. 379-423 and 623-656.

13. Van der Lubbe, J.C.A. (1981), A generalized probabilistic theory on the measurement of certainty and information, Ph. D. Thesis, Dept. E.E., Delft University of Technology, Delft, The Netherlands.

14. Van der Lubbe, J.C.A. (1983), An axiomatic theory of concentration and diversity, National Aerospace Laboratory NLR, Amsterdam, The Netherlands, NLR MP 83015.

Jan C.A. van der Lubbe received his M.Sc. degree in Electrical Engineering from the Delft University of Technology, Delft, The Netherlands, in 1977 and the Ph.D. degree from Delft University of Technology in June 1981.
His Ph.D. dissertation was in statistical information theory.

In 1981 he joined the Remote Sensing Department of the National Aerospace Laboratory (NLR), Amsterdam, The Netherlands. He has been involved in a defintion study for the on board data management for spaceborne radar altimeter under contract with the European Space Agency (ESA). He is currently working in the areas of statistical pattern recognition, texture analysis, image quality assessment and data compression of remotely sensed imagery.