*Magnetic Resonance Imaging*, Vol. 11, pp. 873-887, 1993
Printed in the USA. All rights reserved.

0730-725X/93 $6.00 + .00


● *Original Contribution*

# VIII. MR IMAGE TEXTURE ANALYSIS – AN APPROACH TO TISSUE CHARACTERIZATION

R. A. LERSKI,* K. STRAUGHAN,† L. R. SCHAD,‡ D. BOYCE,† S. BLÜML,‡ AND I. ZUNA‡

*Department of Medical Physics, Ninewells Hospital and Medical School, Dundee DD1 9SY, UK,
†Imperial College of Science, Technology, and Medicine, London, UK, and
‡Institut für Radiologie und Pathophysiologie, Deutsches Krebsforschungszentrum, Heidelberg, Germany

**The role and value of texture analysis in the quantification of medical images is reviewed and the various methods described. The promise in magnetic resonance imaging is discussed and the coordinated research programme being carried out within the framework of the European Economic Community Concerted Action on Tissue Characterization by MRS and MRI is outlined. Tissue characterization of the human brain has been performed by texture analysis of proton relaxation time images using a standard MR whole body imager operating at 1.5 T and the results are presented.**

***Keywords:*** **Magnetic resonance imaging; Tissue characterization; Image texture analysis; $T_1$ and $T_2$ relaxation times.**

## INTRODUCTION

This paper describes some of the work deriving from a collaborative international research project, which forms part of the European Economic Community (EC) Concerted Action on Tissue Characterization by MRI and MRS. The project has addressed the value of texture analytic techniques in deriving robust tissue-characteristic indices from MR images. The work has involved a substantial appraisal of the available techniques and the selection of those techniques most likely to contribute significantly to the tissue characterization problem. Statistical and syntactic methods have been considered, although the statistical methods have been those chosen for the initial work and, therefore, form the dominant theme of this paper. In addition, consideration has been given to the image acquisition process in order to identify optimal and nonoptimal acquisition conditions. Other considerations have included the normalization of images and texture features, optimal classification techniques, and the creation of texture test images.

Given the relative novelty of texture analysis as a postprocessing strategy for MR images, the paper begins with an overview of the principle concepts of texture analysis, including the methods available for classification of the textural features. A review of some key applications of texture analytic techniques in nonclinical and clinical areas is provided to illustrate the power of this class of techniques. The evolution of this project is discussed within the context of the EC Concerted Action and describes the principal strategies and methods adopted. The image analysis tools which have been created during the course of this project and which are available across the three research groups are described. Results from the texture analytic tools are presented both illustratively for test images and for clinical images related to a well-defined clinical trial. The paper ends by exploring the demonstrated value of the texture tools described and the future potential of these and related techniques which are currently under development.

## MOTIVATION

The fundamental objective of any diagnostic imaging investigation is tissue characterization. That is, images are acquired with the express aim of determining if tissues in the chosen region of investigation possess normal or pathological characteristics ("normal" being defined with respect to the typical appearance of healthy tissue for that imaging modality under the specific prevailing image acquisition conditions). Furthermore, given the precise appearance of the relevant

RECEIVED 1/6/93; ACCEPTED 1/6/93.

Address correspondence to R.A. Lerski.

tissue within the image, an attempt is made to classify the pathology as precisely as possible, reference being made to other diagnostic data at hand.

This process, typically performed by the reporting radiologist, requires a complex assessment of the various image features that characterize the appearance of the image under study and the comparison of these observed features with the radiologist's experiential database. Given the inherently subjective nature of many of the judgements associated with this process, the expert radiologist performs this task with a remarkable degree of precision and accuracy.

There exists, however, a significant range of diagnostic problems for which the accuracy of the diagnostic process based on such inspection of image information alone is not acceptable. Furthermore, referring clinicians are increasingly looking for greater specificity in the pathological characterization of tissues from imaging investigations so that they might offer more effective treatment regimes to their patients.

It is this clinical requirement which has provided the motivation for imaging scientists to develop enhanced methodologies, both with regard to the image acquisition system itself and methods for subsequent postprocessing of the image. Activities in the former area have resulted in the emergence of new imaging modalities, such as MRI, and in the dramatic enhancement of the performance of diagnostic imaging systems, very often based on a digital approach. However, it is the development of image processing and image analysis facilities that forms the subject of this paper.

Image processing now constitutes a very large area of study in its own right—the spur to accelerated activity in recent years having been the increasing availability of low-cost workstations of enhanced power. The particular segment of image processing methodology of interest to us here is the use of objective computer-based image analysis to automatically or semi-automatically extract diagnostically significant image features that can be used to distinguish normal from pathological tissue and to further characterize the state of the pathological tissue.

As a starting point, it has been usual for workers in this field to consider those image features which radiologists explicitly or implicitly use in their assessment of the appearance of the specified tissue. Intensity, morphology, and texture are commonly cited as important image features for the radiological assessment of tissue appearance. Image texture is known to be a particularly sensitive feature for the assessment of pathology. The visual assessment of texture is, however, particularly subjective. In addition, it is known that human observers possess a limited sensitivity to textural properties, whereas mathematical texture analytic techniques are significantly more sensitive to textural changes (in addition to being quantitative and, therefore, objective).

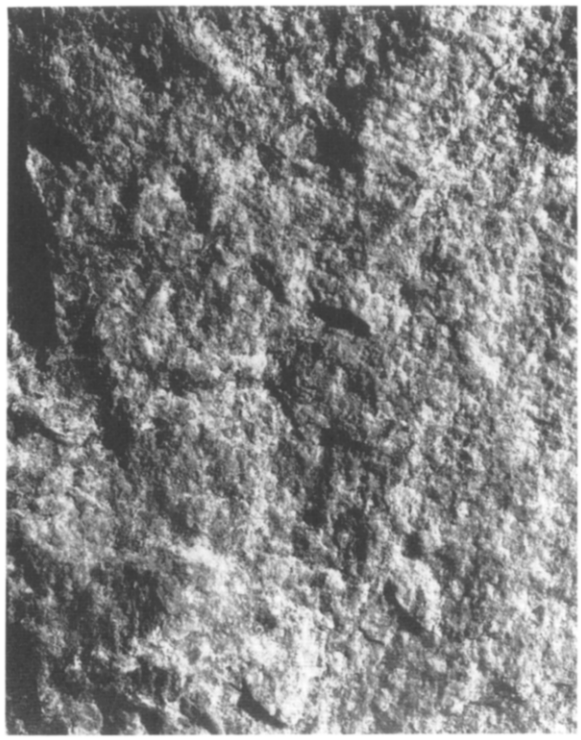
There exists, therefore, considerable motivation for the development and implementation of computer-based texture analysis facilities to derive quantitative textural features for tissues which are of potential value in a tissue characterization task. Such techniques have been used previously with some success in the analysis of clinical ultrasound and CT images (see the section on Implementation of Classification Strategies). In a wider context, texture analytic techniques have been widely developed and implemented in nonclinical imaging applications such as robotic computer vision and remote sensing with considerable success.

## TEXTURE ANALYSIS THEORY

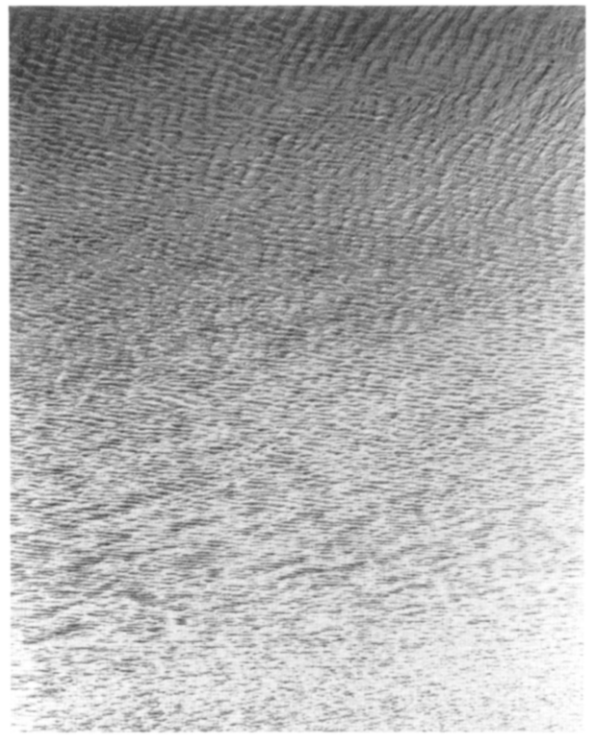
The "texture" of an area of an image is an ill-defined property (Fig. 1). We have an intuitive feeling that it represents the "pattern" of brightness and darkness (grey tones) and that different textures aid us in recognising objects in an image or in discriminating between objects. The visual perception of textures has been the subject of extensive psychological study.[1,2]

*Texture analysis* is a generic name for a series of techniques used for the quantification of spatial variations of grey tones in images. Originally developed for the computer processing of satellite images (to automatically segment such data), a multitude of techniques exist which may be classified into two general categories, namely, statistical and syntactic (or structural). In the first of these, characterization of texture is through the statistics taken from an ensemble of local properties of pixel interrelationships. These methods are best suited to fine textures that contain no obvious regularities of spatial arrangement. In contrast, the structural methods look for an elemental constituent of texture arranged according to some placement rule. Here, the description has to be in terms of both the properties of the structural elements and the form of the placement rules.

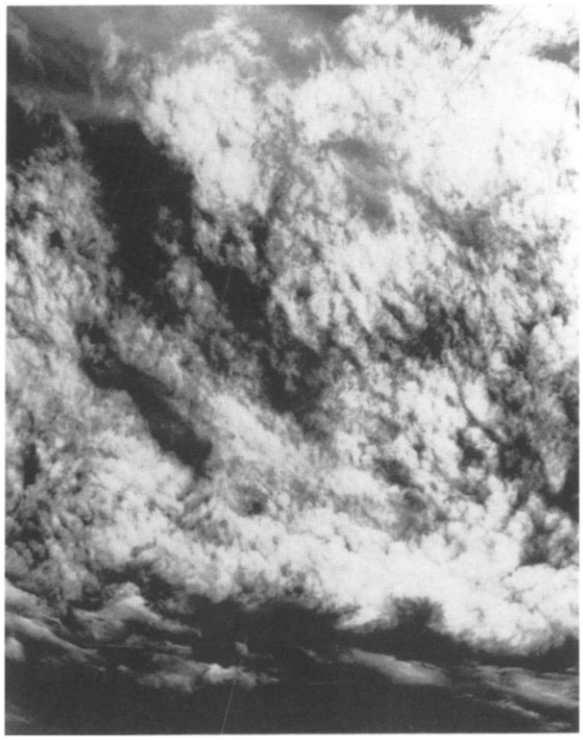
A critical point to stress is that the eye–brain combination is only able to appreciate a limited level of (first and second order) complexity in an image.[3] Texture analysis is able to increase the level of information extracted from the image—this information being inaccessible to human observation. The output is a series of *texture parameters* representing the texture of the region of interest chosen. Whether this information is useful in discrimination has to be determined for each application studied. In medical applications, a training step has to be performed using cases of known pathology to determine the usefulness and to develop a texture-based classifier using the best parameters for the particular discriminatory task. This is discussed in detail in the section on Syntactic Texture Analysis.



(A)



(B)



(C)

Fig. 1. Illustration of a series of textures (A)–(C) to show regular and more random types. (Reprinted from Ref. 33.) Analysis of (A) appears in Table 1 and analysis of (B) appears in Table 2.

*Statistical Methods of Texture Analysis*

Perhaps the most commonly used method of texture analysis is that of grey-tone spatial dependence through a co-occurrence matrix.[4] This is explained in Fig. 2. For a choice of horizontal, vertical, or indeed angled co-occurrence distances, matrices $C(m,n)$ can be constructed whose elements represent the number of times that pixel values $m$ and $n$ occur within the image sepa-

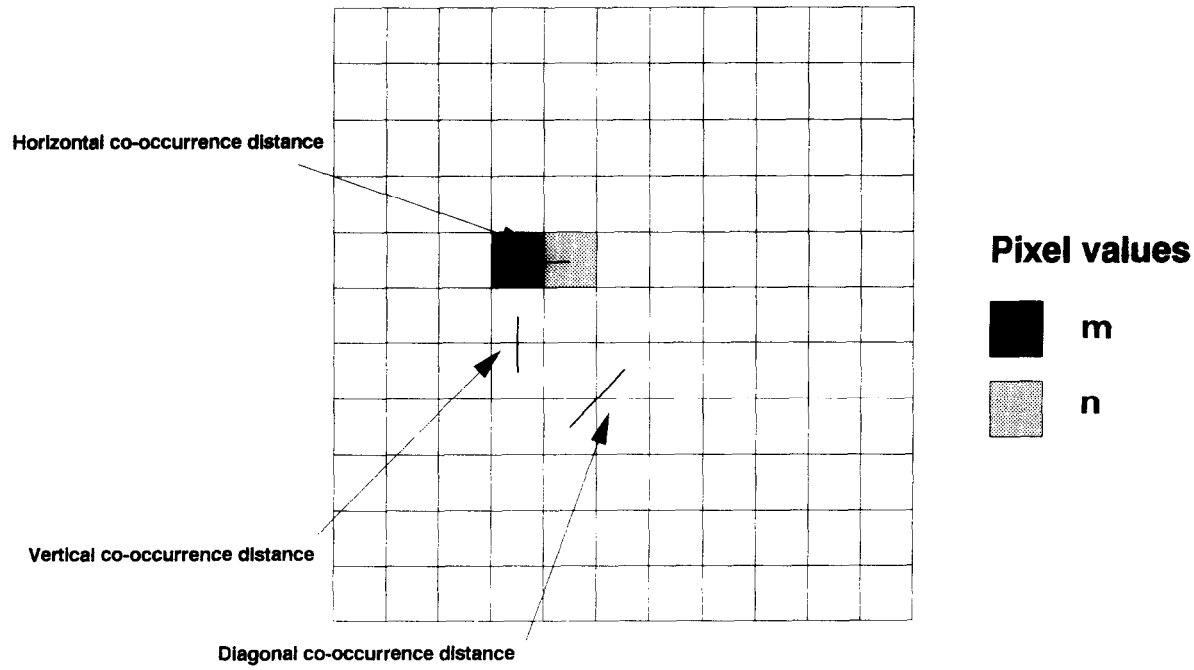



Fig. 2. Grey-tone spatial dependence texture analysis method. An illustration of the construction of the co-occurrence matrix $C(m,n)$ for horizontal, vertical, and diagonal distances.

rated by that distance. Because the matrices contain information on the spatial distribution of pixel values, they are regarded as summarizing the microtexture. From these matrices many parameters may be derived[4,5] that are assumed to represent the texture. The choice of co-occurrence distance is entirely open, but one pixel is the usual selection.

Other commonly applied methods are grey-tone run length analysis,[6] which counts the number of pixels of similar values (within a defined range) that run together in groups, and grey level differencing,[7] which counts the number of times that a chosen difference in pixel values occurs. Weszka et al.[7] compared the methods for a particular application (terrain classification) and mathematical comparison has also been carried out.[8]

In all these methods, it is usually necessary to limit the image grey level values. For example, an 8-bit image would imply a $256 \times 256$ co-occurrence matrix and this would be underpopulated, that is, would have low or zero counts in many of its cells. Often a reduction to, say, 32 levels is used and this may be done either over the whole image or over the selected region of interest.

A crucial factor of all these texture analysis methods is that of the initial normalization of the image data. An attraction of many of the methods is that it is possible to remove sensitivity on the absolute grey-level of the image through a normalization scheme. Several choices exist, for example, zero mean and constant variance or histogram equalization. Also possible is a normalization of, for example, the co-occurrence matrix with respect to the total number of entries. Some reduction of sensitivity will occur, but in many instances there is significant benefit in attempting to remove the dependence of the set-up of the imaging device, at least in terms of gain and contrast. In MRI implementation, it might not be essential to use quantitative relaxation time $T_1$ and $T_2$ images and normalized spin-echo data (see the section on Test Images) may suffice. This approach offers significant savings in terms of data collection time.

*Syntactic Texture Analysis*

As stated previously, the structural or syntactic methods of texture analysis[9–11] are somewhat different in that they attempt to identify fundamental elements of the image or *primitives*, which are connected together to create the entire texture through a syntax or language. The description of the texture is then a definition of the primitives and the syntax that connects them, almost in a linguistic sense (Fig. 3). The various methods that have been used to produce placement rules have been detailed by Matsuyama et al.,[11] varying from adjacency probabilities through graph-like languages, to tree grammars. A combination type of approach has been described by Conners and Harlow,[9] where the inertia measure (see above) of the grey-tone spatial dependence method can be used to

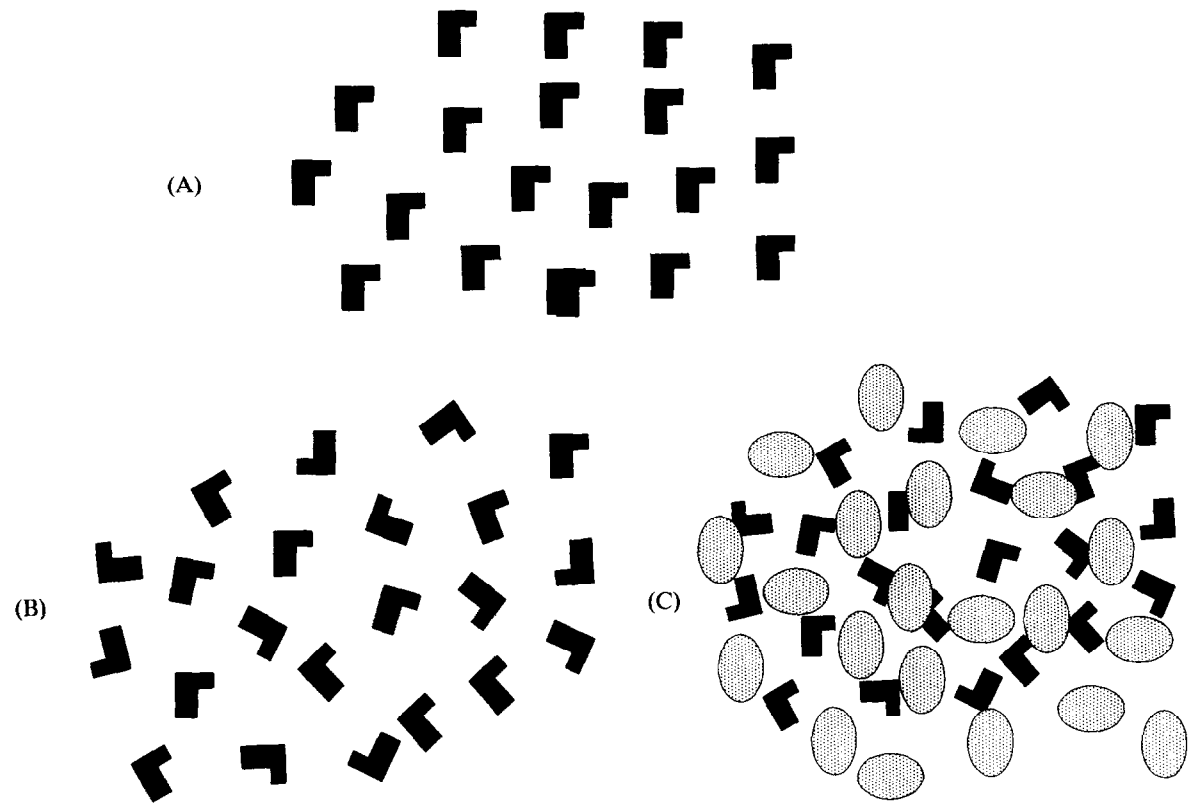



Fig. 3. Textures identified by syntactic methods: (A) a very simple texture containing a single, well-defined primitive; (B) a differing texture with the same primitive, but randomly oriented making it more difficult to identify; and (C) the addition of another primitive oriented in two preferred directions further complicates the procedure.

characterize the placement rules and the unit pattern of regular textures.

These methods are not as highly developed (particularly in the case of medical applications) and are computationally demanding, but may well be very powerful. Generally, artificial intelligence approaches (e.g., PROLOG) are required. They have not been used in the present work.

## IMPLEMENTATION

### *Texture Features*

The 22 features used in this paper are defined in full by Haralick et al.[4] and are summarized in Appendix 1.

### *Classification Strategies*

Usually, the decision problem in human tissue characterization is based on the discrimination between normal tissue and one or more pathologically disordered tissue states. Image texture analysis yields a large number of textural features (see the section on Statistical Methods of Texture Analysis) that are, as a rule, quantitative.

As shown elsewhere,[12] the brightness, microstructure, and macrostructure are the three most diagnostically relevant properties of the examined image. Therefore, the first goal of the classification strategy has to be the selection of the relatively small group of features each of which maps sensitively enough the changes of at least one of these properties.

In the second step, the optimum classifier in terms of dependency on the statistical distribution of the selected features and on the defined diagnostic classes has to be designed. More precisely, the classification model for pattern recognition contains in general three parts, a transducer, a feature extractor, and a classifier.[13] The transducer senses the input and converts it into a diagnostic image. The feature extractor derives presumably relevant information from this image and the classifier uses this information to assign the input data to one of a finite number of diagnostic classes.

We are concerned with both feature extraction and classification. Feature extraction is much more problem-dependent than classification. The problem of classification is basically one of partitioning the feature space into regions, one region for each category. Ideally, this partitioning would be arranged so that none of the decisions is incorrect or the probability of error

would be minimized. In this case, the problem of classification becomes a problem in statistical decision theory.

*Feature extraction.* Before looking into the discriminant power of each particular texture feature, the grade of their mutual "interdependence" must be analyzed. Using correlation analysis, or more often the powerful technique of factor analysis, the groups of strongly mutually dependent features have to be found to select the subgroup that are approximately independent (orthogonal).

After that, the most promising method, step-wise discriminant analysis with the Mahalanobis or Euclidean distance measure,[14] is used to find the group of features yielding the best discrimination results for the given diagnostic problem. In this case, an analysis of variance $F$-test is used to measure the respective discriminatory power. To avoid the "statistical trap" of overestimation,[15] the final number of selected features must always be smaller than a third of the number of images in the smallest diagnostic class.

*Classification.* Depending on the type of statistical distribution of the selected feature vector, linear or quadratic discriminant analysis and the kernel or k-nearest-neighbor methods are the most promising methods for classification.[16] When the distribution within each group is assumed to be multivariate normal, a parametric method can be used to develop a discriminant function. It is determined by a measure of generalized squared distance between the centres of all diagnostic groups. The classification criterion is based on either the individual within-group covariance matrices (yielding a quadratic function) or the pooled covariance matrix (yielding a linear function).

When no assumption can be made about the distribution within each group, or when the distribution is assumed to be different from a multivariate normal, the nonparametric kernel or k-nearest-neighbor method can be used to estimate the group-specific densities. With a k-nearest-neighbor method, the pooled covariance matrix is used to calculate the distances between the diagnostic classes. With a kernel method, both individual and pooled matrices can be used. With the estimated group-specific densities, the most probable group membership for each class can be evaluated.

## APPLICATIONS

*Nonclinical*

The original implementation of texture analysis technique was for the automatic computer segmentation of aerial photographs, for example, LANDSAT images. This extended to military applications and geological studies with the intention of dealing with large amounts of image data in a rapid manner. Many publications have dealt with training tests of the ability of various texture analysis methods to classify such data.[7] Also common have been comparative studies, in a mathematical sense, of the discrimination ability of different methods.[8]

The general conclusion of this work is that standard methods, such as grey-tone spatial dependence through the co-occurrence matrix, can automatically segment aerial photographic data with a high degree of reliability.

*Clinical*

In certain medical imaging techniques, for example, ultrasound, the *visual* texture is recognized as conveying diagnostic information. In other methods, for example, X-ray CAT or MRI, its use is not yet established. Medical applications of computerized texture analysis date back to the early 1970s, when it became possible to digitize X rays for computer processing. The explosion in the use of digital techniques has led to the possibility of utilizing texture analysis for most medical imaging modalities. Typical work has been carried out on chest X-rays,[17-20] ultrasound,[21-25] and X-ray CAT[26] (also Lerski and Straughan, in preparation).

To date, the studies carried out, although having demonstrated a good measure of success, have not yet been accepted into routine clinical practice. In ultrasound, normal and amyloid myocardial structures have been distinguished,[21] prostatic carcinoma and prostatic hypertrophy separated,[22] placental condition quantified,[24] and hepatitis differentiated from normal liver—a result that cannot be achieved by conventional ultrasonography.[23,27] In chest X rays, a close correspondence was found between the International Labor Office classification categories for pneumoconioses and texture measures from computer analysis.[28] All these successes demonstrate an ability to extract more objective diagnostic information from the images than can be done by a human observer.

## THE EC COLLABORATION—EVOLUTION, STRATEGIES, AND METHODS

The EC Concerted Action (CA) on Tissue Characterization by MRI and MRS began in 1984 with the clear mission to explore the use of parametric data derived from in vivo MRI and MRS studies as tissue characteristic indices. Much of the early work of the CA revolved around the development of standardization and calibration techniques to enable the comparison of parametric image data (most notably the spin relaxation times $T_1$ and $T_2$) derived from different MRI systems. Subsequent to this work, a significant activity of the

CA was the development of methods for optimal relaxation time calculation and comparison and the establishment of a pan-European database of in vivo $T_1$ and $T_2$ values. Some of this work has been previously described in this journal.[29]

As part of the evolution of the CA, parametric indices other than the spin relaxation times were explored as candidates for tissue characterization. Textural features were of especial interest due to the previous experience of several of the participants of the CA in this field, albeit using different imaging modalities. A specialist working party was constituted, which met first in 1987, to explore the potential role of texture analytic techniques and to formulate a research strategy. The consensus of that meeting[30] was that textural image features possessed significant potential as image features which might convey tissue characteristic information.

The strategy formulated at that meeting[31] revolved around the exchange of image data between participating sites so that comparative texture features could be derived from identical data sets using the various analytic tools developed at individual centres. While individual centres were free to classify their own features, all derived texture features were to be submitted to one centre for classification using an extensive standard suite of classification tools. Although various protocols were conceived, it was decided that the specification of rigid image acquisition protocols would be counterproductive at that stage because too little was known about the sensitivity of the various texture features to the exact form of the image contrast function. Similarly, while agreeing that image normalization was an essential prerequisite to the texture analysis (especially given the desire to perform comparative studies involving a number of very different MR systems), the precise specification of the normalization procedure was thought to require detailed investigation involving the comparison of texture features from un-normalized and normalized images.

Texture "test objects" were thought to be necessary for the objective comparison and calibration of the numerous texture analytic tools to be studied. The enormous difficulty associated with producing clinically realistic textures which were MR compatible, reproducible, and stable was widely acknowledged. Some ideas were exchanged as to potential approaches to this problem, but further progress was deemed to require a separate research activity. In the absence of such test objects, the use of synthetic texture images was adopted. Such images can be derived either using mathematical models or using photographic methods. Both avenues would be explored to derive a set of appropriate texture test images.

Subsequent to the 1987 workshop, three Centres (DKFZ, Heidelberg; Imperial College, London; University of Dundee) have formed a core collaborative program to carry forward the proposals of the CA Working Group. Clinical and test images have been exchanged between the centres. In addition, software implementation of some of the texture analytic tools used within the project have been exchanged such that identical tools are available at all three principal centres. While feature classifiers are available at all three sites, the extensive resource of feature classifiers at Heidelberg has been used for common analysis. More recently, the centres have been contemplating the development of a Unix-based software utility to provide a common and portable image processing platform on which to base a more widespread pan-European clinical trial.

## TEST IMAGES

At the 1987 CA workshop, a number of strategies were considered for the development of mathematical synthetic texture images. Subsequent work has shown that, on the whole, these synthetic images were inappropriate as they were based on rather too simplistic a model. More recently, other mathematical models have been demonstrated to have value in other applications.[32] Such synthetic images are currently being explored within the framework of this project.

Given the disappointing characteristics of the initial mathematical synthetic images and pending the implementation of the new proposals, photographically derived texture images have been utilized to provide test images. In common with the leading workers on texture, we have adopted the set of textures published by Brodatz.[33] From this extensive set, we have selected a subset of textures that possess characteristics which are broadly compatible with those to be encountered in clinical images. The images have been subsequently digitized with an 8-bit dynamic range at a spatial resolution of 512 × 512. The selected images have been distributed to the complete CA working group.

Results from two of the selected images have been chosen for presentation here. The two images are illustrated in Figs. 1A and 1B.

These test images have been analyzed using statistical texture analytic tools, with the Haralick co-occurrence features[4] being used as a key element in the analysis (see Appendix 1 for parameter definitions). The illustrative results are given in Tables 1 and 2. They show that such test images have appreciable angular dependence in their texture, with both the values of texture parameters at particular angles and the ranges over the angles chosen having possible discriminatory value.

Table 1. Test image 1, ROI = top left hand corner

| Parameter | Co-occurrence direction 0° | 45° | 90° | 135° |
|---|---|---|---|---|
| F1 | 0.0021 | 0.0018 | 0.0023 | 0.0020 |
| F2 | 18.499 | 28.132 | 17.474 | 20.755 |
| F3 | 0.9729 | 0.9542 | 0.9745 | 0.9683 |
| F4 | 222.82 | 220.96 | 221.46 | 220.98 |
| F5 | 0.3073 | 0.2564 | 0.3205 | 0.2970 |
| F6 | 177.89 | 177.94 | 178.00 | 177.95 |
| F7 | 879.19 | 863.62 | 875.55 | 869.89 |
| F8 | 4.6757 | 4.6724 | 4.6801 | 4.6770 |
| F9 | 6.5173 | 6.7355 | 6.4742 | 6.5746 |
| F10 | 4.8030 | 8.3662 | 4.6547 | 5.7852 |
| F11 | 2.0589 | 2.2885 | 2.0249 | 2.1220 |
| F12 | 0.2524 | 0.2525 | 0.2524 | 0.2525 |
| F13 | 0.9399 | 0.9064 | 0.9449 | 0.9322 |
| F14 | 0.4627 | 0.5075 | 0.5345 | 0.7071 |

Test image 1 is illustrated in Fig. 1A.

The spatial position of the region of interest is also clearly of strong influence on the results obtained.

Further, Fig. 4A illustrates the variation of two of the texture features over a wide range of eight regions of interest in the first test image. Figure 4B illustrates the variation of the same features in four regions of interest in the three test images (Figs. 1A–1C)—a discrimination of these images simply on the basis of these results would be possible.

These brief results confirm the great value of such test images in validating the implementation of the texture software at different sites (due to the sensitivity of the analysis) and are forming the basis of continuing work within the CA.

Table 2. Test image 2, ROI = bottom right hand corner

| Parameter | Co-occurrence direction 0° | 45° | 90° | 135° |
|---|---|---|---|---|
| F1 | 0.0025 | 0.0026 | 0.0038 | 0.0025 |
| F2 | 35.664 | 38.029 | 18.217 | 41.890 |
| F3 | 0.7417 | 0.7180 | 0.8834 | 0.6883 |
| F4 | 51.082 | 51.038 | 51.020 | 51.036 |
| F5 | 0.2065 | 0.2259 | 0.3131 | 0.2176 |
| F6 | 147.33 | 147.32 | 147.34 | 147.32 |
| F7 | 177.94 | 175.36 | 192.17 | 172.33 |
| F8 | 3.9956 | 3.9889 | 4.0304 | 3.9812 |
| F9 | 6.2654 | 6.3021 | 5.9195 | 6.3385 |
| F10 | 9.2918 | 11.812 | 4.8683 | 13.027 |
| F11 | 2.4129 | 2.4575 | 2.0461 | 2.5066 |
| F12 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| F13 | 0.5814 | 0.5488 | 0.7900 | 0.5147 |
| F14 | 0.3710 | 0.3462 | 0.5000 | 0.3806 |

Test image 2 is illustrated in Fig. 1B.

## CLINICAL RESULTS

### *Introduction*

As shown in previous work in ultrasonic tissue characterization,[12] several parameters, derived from the statistical distribution of grey levels in the image, showed a very close correlation to the morphological, biochemical, and histological properties of the underlying tissue. These parameters have been applied for the successful classification and discrimination of pathological tissues in organs such as liver, thyroid, breast, and kidney. These results encouraged the application of pattern recognition methods to tissue characterization in MRI.

A small, prospective, clinical study of the textural characteristics of MR images of selected brain tumors was, therefore, undertaken. The aim of this study was to describe the texture of different tissue types by use of image texture features from calculated $T_1$ and $T_2$ parameter images, as well as from the proton density image, for example, by using parameters from the first and second order grey level and gradient statistics. This work is briefly described here as an illustration of the potential diagnostic utility of texture methods of image analysis.

### *Patients, Examinations, and Methods*

In this prospective study, 12 patients with brain tumors were examined. In 10 of these patients the tumor histology was known: 5 patients had glioblastomas and 5 patients had tumors arising from metastases. Most of the patients suffered from brain oedema.

A total of 113 regions of interest (ROI) were defined by two experienced radiologists in the $T_1$ or $T_2$ parameter images of the selected slices. Of these, only regions with size greater than 200 pixels were selected. This left 78 remaining regions that fell into the following five tissue classes: tumor (8 ROIs), oedema (11 ROIs), liquor (12 ROIs), white matter (24 ROIs), and grey matter (23 ROIs).

Delineation of the tumor was best in the post-contrast images, that is, the tumor ROI was defined by regions of Gd-DTPA enhancement in post-contrast images and then transferred to the pre-contrast images.

$T_1$ and $T_2$ measurements were performed on a 64-MHz MAGNETOM (Siemens, Erlangen, Germany) whole body imager. For diagnostic imaging, $T_1$ and $T_2$ weighted SE images in transversal or coronal orientations were performed (TR/TE 500/325 and TR/TE 1600/35,70, respectively) and the slices with the best visualization of the brain tumor were chosen for the

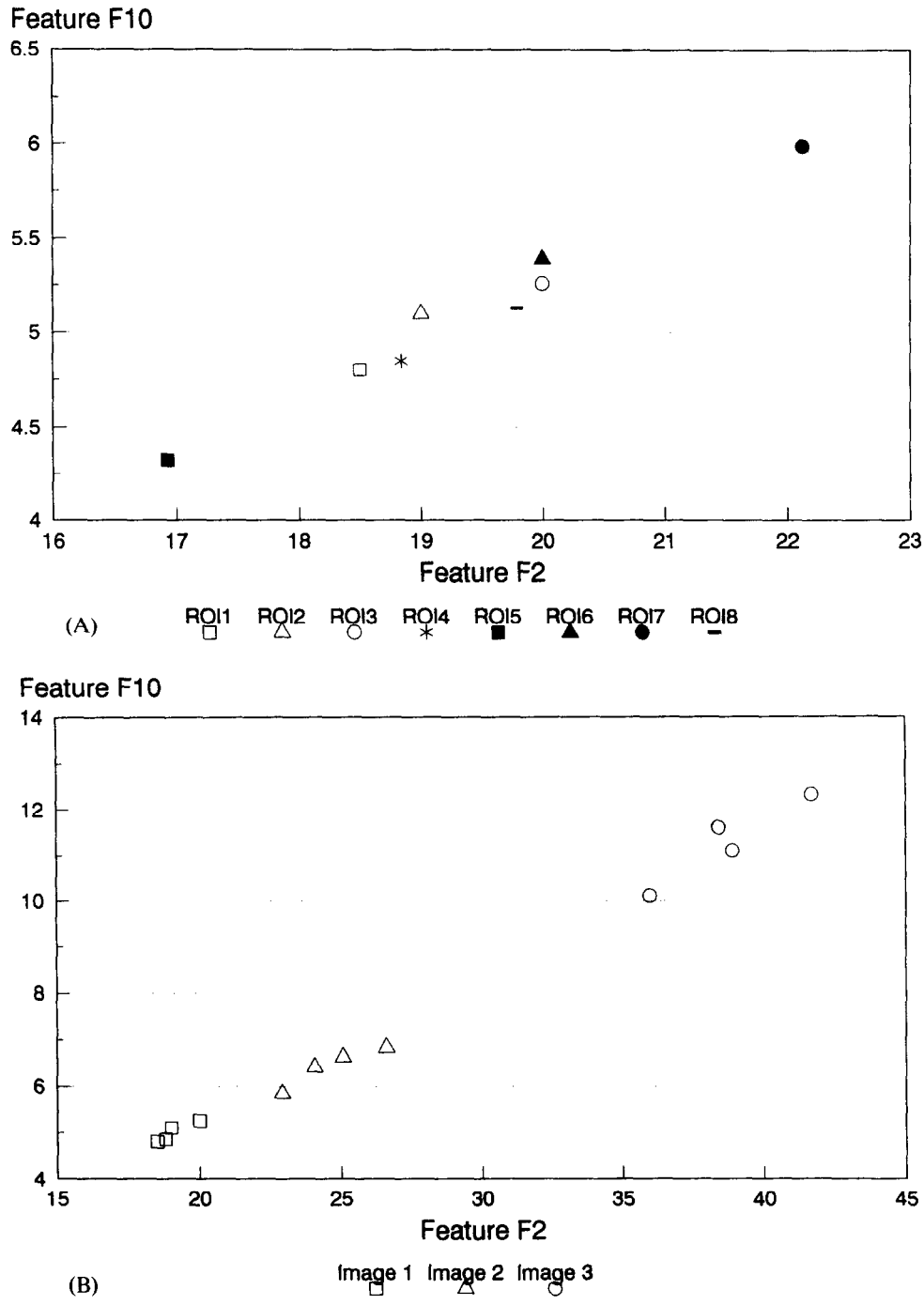



Fig. 4. Scatter diagrams of two texture features (A) in eight regions of interest in test image 1 and (B) from four regions of interest in the three test images (Figs. 1A-1C).

calculation of $T_1$, $T_2$, and proton density parameter images.

To examine the ability of statistical pattern recognition methods for MR tissue characterization, first and second order image texture parameters were computed in the ROIs from the parameter images to describe quantitatively the brightness, micro-, and macrotexture of normal and pathological human brain tissues.

The following describes the most important groups of statistical parameters which characterize the first and second order image texture and their figurative interpretations are given briefly. Exact mathematical defi-

nitions of all image texture features used in this study can be found in the original paper of Haralick et al.[4]

The mean grey level (MGL) gives the average value of grey levels belonging to all pixels in the ROI, for example, the average brightness of the ROI. By having an approximate gaussian distribution of grey levels in the ROI, the variance of grey levels (VGL) characterizes the shape of the grey level histogram. In the standard tissue characterization approach, only the MGL is used as a quantitative measure of brightness for the classification of tissues.

The gradient parameters describe the local distribution of grey level differences, for example, the microtexture or coarseness of the ROI. The parameters used are mean gradient (MGR) and variance of gradients (VGR) in the ROI.

The second order parameters originate from the grey level co-occurrence matrix. From each co-occurrence matrix four parameters were calculated: contrast (CON), angular second moment (ASM), entropy (ENT), and correlation (COR). Contrast is sensitive to large local grey level differences. Angular second moment and entropy characterize the homogeneity of the matrix entries. Correlation detects the presence of spatial homogeneous subregions in the ROI and characterizes, therefore, the macrotexture of the ROI.

*Feature Selection and Classification*

All features obtained by image texture analysis were statistically analyzed.[34] Features that demonstrated strongly skewed distributions were rejected from the study. Using a correlation analysis, strongly dependent features were also rejected. A step-wise discriminant analysis was used to select the most discriminating parameters at each hierarchical level. Because the patients in this study were preselected, equal prior probabilities were used. For the measurement of discriminatory power an analysis of variance *F*-Test was applied.

The classification method used in this study was based on four two-class discriminant tests in a four-layer hierarchical decision tree. In the first classification layer, a decision was made whether the tissue structure is liquor. After this test, a decision was made to separate the class of white matter from the other classes. In the third decision, the class of grey matter was separated from the remaining classes tumor and oedema. Finally, the fourth discriminant test separated the class tumor from the oedema tissue pattern.

*Results*

*First layer: Separation of liquor regions.* In the first layer, the group was discriminated which contained liquor regions (12 ROIs) from all other regions (66 ROIs). There was no difficulty in distinguishing between these two groups by using the MGL from the $T_2$ image (Table 3). The accuracy, defined as the ratio of the correctly classified regions to all regions being classified, was 100%.

Table 3. The best features for discrimination between liquor and common group of white and grey matter, oedema, and tumour by both standard and optimum approaches

| Approach | Best features | Accuracy |
|---|---|---|
| Standard | MGL-$T_2$ | 100% |
| Optimum | MGL-$T_2$ | 100% |

MGL = mean grey level.

*Second layer: Separation of white matter regions.* In the second layer, discrimination was carried out between the group that contained white matter regions (24 ROIs) and the group of all other regions without liquor (42 ROIs). Using the best combination of brightness features (standard approach), we obtained an overall accuracy of 97% with two regions being wrongly classified as white matter. By combining the MGL from the $T_1$ image and microtexture measure entropy from the $T_2$ image, an accuracy of 100% was reached (Table 4).

*Third layer: Separation of grey matter regions.* In the third layer, discrimination was carried out between the group that contained grey matter regions (23 ROIs) and the group having both tumor and oedema (19 ROIs). The best combination of brightness features yielded an overall accuracy of 95%; two regions were misclassified. The optimum feature vector containing the MGL from the $T_2$ image combined with the MGL from the proton density image and entropy from the $T_1$ image yielded an overall accuracy of 100% with one wrongly classified grey matter region (Table 5).

Table 4. The best features for discrimination between white matter and common group of grey matter, oedema, and tumour by both standard and optimum approaches

| Approach | Best features | Accuracy |
|---|---|---|
| Standard | MGL-$T_1$, MGL-$T_2$ | 100% |
| Optimum | MGL-$T_1$, ENT-$T_2$ | 100% |

MGL = mean grey level; ENT = entropy.

Table 5. The best features for discrimination between grey matter and common group of oedema and tumour by both standard and optimum approaches

| Approach | Best features | Accuracy |
|---|---|---|
| Standard | MGL-$T_1$, MGL-$T_2$ | 95% |
| Optimum | MGL-$T_2$, MGL-RH, ENT-$T_1$ | 100% |

MGL = mean grey level; RH = proton density image; ENT = entropy.

Table 6. The best features for discrimination between oedema and tumour by both standard and optimum approaches

| Approach | Best features | Accuracy |
|---|---|---|
| Standard | MGL-$T_1$, MGL-$T_2$, MGL-RH | 74% |
| Optimum | MGR-$T_2$, COR-$T_1$, RLDIST-RH | 95% |

MGL = mean grey level; RH = proton density image; MGR = mean gradient; COR = correlation; RLDIST = runlength MGL distribution.

*Fourth layer: Differentiation between tumors and oedematous regions.* In the last layer, discrimination was carried out between the group that contained tumor regions (8 ROIs) and the group with oedematous regions (11 ROIs). The discrimination using just the brightness features was completely misleading. The feature vector containing MGL values from all three parameter images led to an overall accuracy of 74%, with five misclassified regions. However, by using the optimum feature combination of mean gradient from the $T_2$ parameter image with correlation from the $T_1$ parameter image and the runlength MGL distribution from the proton density image, an accuracy of 95% was achieved with one misclassified oedema region (Table 6).

The separation at the third and fourth level can be done in one step by using only two parameters, as illustrated in Fig. 5. The three groups of tissue regions, corresponding to grey matter (●), oedema (■), and tumor (▲) can be separated by using the MGL and the MGR from the $T_2$ image. As can be seen, with this technique there were two misclassifications, one in which grey matter was interpreted as tumor, and one in which a hyper-nephroma-metastasis was wrongly classified as grey matter.

In the hope of describing the response of a texture parameter on different tissue structures as clearly as possible, the discrimination results are presented as shown in the Fig. 6. The cluster of grey matter regions (●) show low values of brightness (MGL) and medium

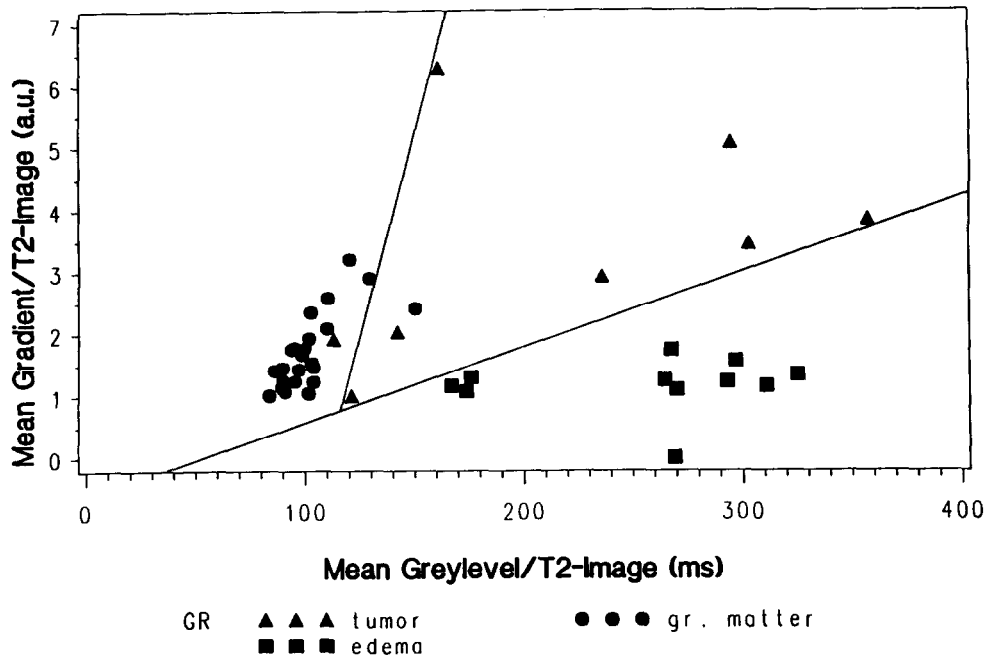



Fig. 5. Best discrimination between three groups of tissue regions which correspond to grey matter (● $n = 23$), oedema (■ $n = 11$), and tumour (▲ $n = 8$) by using the mean grey level (MGL) and mean gradient (MGR) of the $T_2$ image.

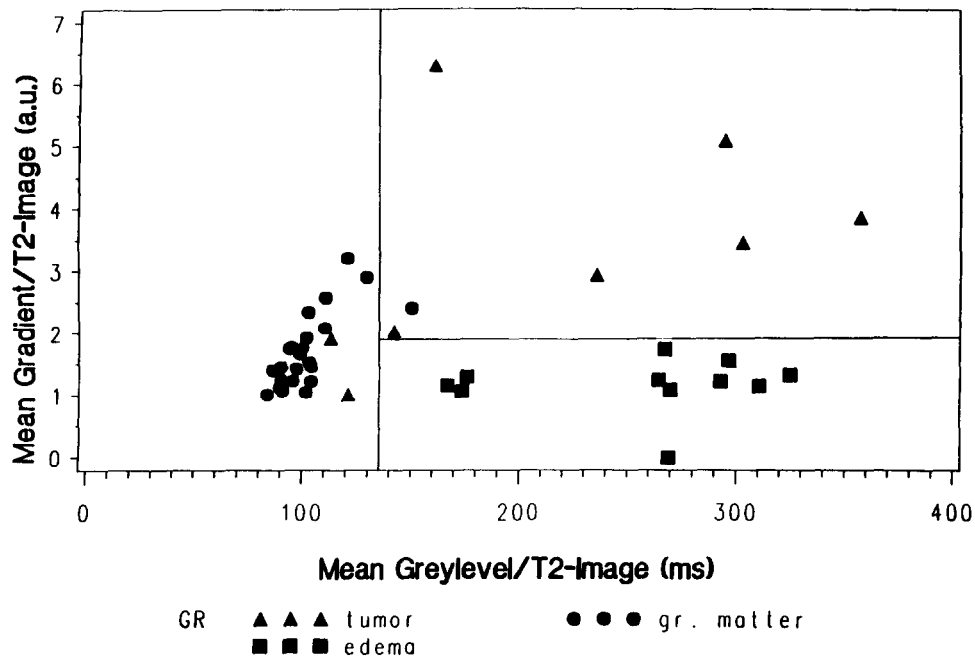



Fig. 6. Vertical and horizontal separation of three groups of tissue regions which correspond to grey matter (● $n = 23$), oedema (■ $n = 11$), and tumour (▲ $n = 8$) by using the mean grey level (MGL) and mean gradient (MGR) of the $T_2$ image.

values of texture-coarseness (MGR) in feature space, for example, the grey matter regions are usually darker than oedematous and tumor regions but there is no significant difference in the coarseness between grey matter and pathological regions. The separation of the grey matter regions can therefore only be done by the use of the MGL of the $T_2$ parameter image (vertical separation line).

Further, the MGL for both oedematous (■) and tumor tissue (▲) clusters is approximately the same but is substantially higher than the MGL for the grey matter cluster. Thus, it is very difficult to differentiate between tumor tissue and oedema using a standard approach. However, the mean coarseness of the tumor tissue cluster is significantly higher than the mean coarseness of the oedematous cluster and can be used for discrimination. Thus, the separation between oedema and tumor is therefore possible by usage of MGR of the $T_2$ parameter image (horizontal separation line). (The best discrimination of the three groups has been shown in Fig. 5.)

In conclusion, the results of this preliminary study suggest that statistical image texture analysis of MR images of human brain is of clinical importance in the discrimination between brain tumor and oedema, but needs further verification by carrying out a well-defined clinical study with at least 20 patients in each diagnostic group.

## DISCUSSION – FUTURE POTENTIAL

The benefit of texture analysis is the extraction of extra discriminatory information from digital medical images, in particular, features that are inaccessible to human observation (the higher order statistics). The overall information may be used for tissue characterization and there is widespread evidence from several modalities that this can be achieved fairly successfully. The possible reason that widespread application has not been made, is that training of the texture classifier is required using cases of known pathology and, for best results, this training dataset should be large. There are also questions to be answered regarding the robustness of the texture measures to different instrumentation and the best schemes of normalization to be employed.

Nevertheless, as has been demonstrated here, there is exciting potential in the application of texture analysis, and in MR brain imaging, significantly enhanced information can be extracted from the image data. Further improvements may be expected from refinements in standardization methodologies and from the introduction of syntactic techniques, specifically designed for medical imaging modalities.

A substantial collaborative research programme designed to further develop, enhance, and evaluate these techniques is currently ongoing within the framework of the EC Concerted Action.

## APPENDIX 1: DEFINITIONS OF TEXTURE PARAMETERS

| | |
|---|---|
| Region of Interest | $R$ |
| Pixels | $N$ |
| Grey Levels | $g(i,j)$ |
| No. of grey levels in image | $N_g$ |

### (a) First Order

Mean Grey Level (MGL)

$$\frac{1}{N}\sum_{ij} g(i,j)$$

Variance of Grey Levels (VGL)

$$\frac{1}{N}\sum_{ij} (g(i,j) - MGL)^2$$

For gradient calculations a neighborhood of pixels is used viz.

| | | | | |
|---|---|---|---|---|
| A | B | C | D | E |
| F | G | H | I | J |
| K | L | M | N | O |
| P | Q | R | S | T |
| U | V | W | X | Y |

Absolute gradient value [ABSV$(i,j)$]

$$\sqrt{(W-C)^2 + (O-K)^2}$$

Mean gradient value (MGR)

$$\frac{1}{N}\sum_{ij} ABSV(i,j)$$

Variance of gradients (VGR)

$$\frac{1}{N}\sum_{ij} (ABSV(i,j) - MGR)^2$$

### (b) Second Order

From co-occurrence matrix

$$p(i,j)$$

$p_x(i)$ $i$th entry in the marginal probability matrix obtained by summing the rows of $p(i,j)$

$$\sum_{j=1}^{N_g} p(i,j)$$

similarly, $p_y(j)$

$$\sum_{i=1}^{N_g} p(i,j)$$

$p_{x+y}(k)$

$$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)$$

$$i + j = k$$

$$k = 2,3,\ldots,2N_g$$

$p_{x-y}(k)$

$$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)$$

$$|i - j| = k$$

$$k = 0,1,\ldots,N_g - 1$$

then define parameters

F1 (ASM)

$$\sum_{ij} p(i,j)$$

F2 (CON)

$$\sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \right\}$$
$$|i-j| = n$$

F3 (COR)

$$\sum_i \sum_j \frac{ijp(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$$

where

$\mu_x$, $\mu_y$, $\sigma_x$, $\sigma_y$ are the means, standard deviations of $p_x$, $p_y$

F4

$$\sum_i \sum_j (i-\mu)^2 p(i,j)$$

F5

$$\sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j)$$

F6

$$\sum_{i=2}^{2N_g} i p_{x+y}(i)$$

F7

$$\sum_{i=2}^{2N_g} (i-f_8)^2 p_{x+y}(i)$$

F8

$$-\sum_{i=2}^{2N_g} p_{x+y}(i) . \log\{p_{x+y}(i)\}$$

F9

$$-\sum_i \sum_j p(i,j) . \log\{p(i,j)\}$$

F10

variance of $p_{x-y}$

F11

$$-\sum_{i=0}^{N_g-1} p_{x-y}(i) . \log\{p_{x-y}(i)\}$$

F12

$$\frac{HXY - HXY1}{\max\{HX, HY\}}$$

F13

$$(1 - \exp\{-2.0(HXY2 - HXY)\})^{0.5}$$

where

$$HXY = -\sum_i \sum_j p(i,j) \log\{p(i,j)\}$$

$HX$ and $HY$ are the entropies of $p_x$ and $p_y$ and

$$HXY1 = -\sum_i \sum_j p(i,j) \log\{p_x(i)p_y(j)\}$$

and

$$HXY2 = -\sum_i \sum_j p_x(i)p_y(j) \log\{p_x(i)p_y(j)\}$$

F14

(Second largest eigenvalue of Q)$^{0.5}$

where

$$Q(i,j) = \sum_k \frac{p(i,k) . p(j,k)}{p_x(i)p_y(k)}$$

## REFERENCES


1. Callaghan, T.C. Interference and dominance in texture segregation: Hue, geometric form and line of orientation. *Percept. Psychophys.* 46:299; 1989.
2. Pickett, R.M. The perception of a visual texture. *J. Exp. Psychol.* 68:13; 1964.
3. Julesz, B. Visual pattern discrimination. *IRE Trans. Info. Theory* 8:84-92; 1962.
4. Haralick, R.M.; Shanmugam, K.; Dinstein, J. Textural features for image classification. *IEEE Trans. Syst. Man. Cybern.* 6:610-62; 1973.
5. Räth, U.; Schlaps, D.; Limberg, B.; Zuna, I.; Lorenz, A.; van Kaick, G.; Lorenz, W.; Kommerell, B. Diagnostic accuracy of computerized B-scan texture analysis and conventional ultrasonography in diffuse parenchymal and malignant liver disease. *J. Clin. Ultrasound* 3:87-99; 1985.
6. Galloway, M.M. Texture analysis using grey level run lengths. *Comput. Graph. Image Process.* 4:172-179; 1975.
7. Weszka, J.S.; Dyer, C.R.; Rosenfeld, A. A comparative study of texture measures for terrain classification. *IEEE Trans. Syst. Man. Cybern.* 6:269-285; 1976.
8. Conners, R.W.; Harlow, C.A. A theoretical comparison of texture algorithms. *IEEE Trans. Pattern Anal. Machine Intel.* 2:204-222; 1980.
9. Conners, R.W.; Harlow, C.A. Towards a structural textural analyser based on statistical methods. *Comput. Vision, Graph. Image Process.* 12:224-256; 1980.
10. Haralick, R.M. Statistical and structural approaches to texture. *Proc. IEEE* 67:786-804; 1979.
11. Matsuyama, T.; Saburi, K.; Nagao, M. A structural analyser for regularly arranged textures. *Comput. Vision, Graph. Image Process.* 18:259-278; 1982.
12. Zuna, I. Computerized ultrasonic tissue characterization: Methods and clinical use. In: H.U. Lemke, et al. (Eds). Computer Assisted Radiology – CAR'87. Berlin:Springer; 1987:155-163.
13. Duda, R.O.; Hart, P.E. Pattern Classification and Scene Analysis. New York: Wiley-Interscience; 1973.
14. Lachenbruch, P.A. Discriminant Analysis. London: Hafner Press; 1975.
15. Foley, D.H. Considerations of sample and feature size. *IEEE Trans. Inform Theory.* 18:618-626; 1972.
16. Parzen, E. On estimation of a probability density function and mode. *Ann. Math. Stat.* 8:247-263; 1962.
17. Chien, Y.P.; Fu, K.S. Recognition of x-ray picture patterns. *IEEE Trans. Syst. Cybern.* 4:145; 1974.
18. Desaga, J.F.; Dengler, J.; et al. Film digital and texture analysis for digital classification of pulmonary spot opacities. *Rontgenblatter* 41:147; 1988.
19. Powell, G.F.; Doi, K.; Katsuragawa, L. Localization of inter-rib spaces for lung texture analysis and computer aided diagnosis in digital chest images. *Med. Phys.* 15:581; 1988.
20. Sutton, R.N.; Hall, E.L. Texture measures for automatic classification of pulmonary disease. *IEEE Trans. Comput.* 21:667; 1972.
21. Chandrasekran, K.; Aylward, P.E.; Fleagle, S.R.; et al. Feasibility of identifying amyloid and hypertrophic cardiomyopathy with the use of computerized quantitative texture analysis of clinical echocardiographic data. *J. Am. Coll. Cardiol.* 13:832; 1989.
22. Kratzik, C.; Schuster, E.; Hainz, A.; Kuber, W.; Lunglmayr, G. Texture analysis – A new method for discriminating prostatic carcinoma from prostatic hypertrophy. *Urol. Res.* 16:395; 1988.
23. Lerski, R.A.; Barnett, E.; Morley, P.; et al. Computer analysis of ultrasonic signals in diffuse liver disease. *Ultrasound Med. Biol.* 5:341; 1979.
24. Morris, D.T. An evaluation of the use of texture measurements for the tissue characterisation of ultrasonic images of in-vivo human placenta. *Ultrasound Med. Biol.* 14:387; 1988.
25. Raeth, U.; Schlaps, D.; Limberg, B.; et al. Diagnostic accuracy of computerized B-scan texture analysis and conventional ultrasonography in diffuse parenchymal and malignant disease. *J. Clin. Ultrasound* 13:87; 1985.
26. Coleman, A.J.; Tonge, K.A.; Rankin, S.C. The power spectral density as a texture measure in computed tomographic scans of the liver. *Br. J. Radiol.* 55:601; 1982.
27. Schuster, E.; Knoflach, P.; Grabner, G. Local texture analysis: An approach to differentiating liver tissue objectively. *J. Clin. Ultrasound* 16:453; 1988.
28. Katsuragawa, S.; Doi, K.; Macmahon, H. Quantitative computer aided analysis of lung texture in chest radiographs. *Radiographics* 10:257; 1990.
29. Concerted research project of the European Economic Community. Identification and characterization of Biological Tissues by NMR. *Magn. Reson. Imaging* 6:173-222; 1988.
30. Lerski, R.A. Working group on texture analysis in NMR imaging. *Eurospin Quarterly* 12:48-89; 1987.
31. Lerski, R.A.; Straughan, K. Report from the Working Group on Texture Analysis in NMR imaging. *Eurospin Quarterly* 12:79-89; 1987.
32. de Kedrel, T.; Constantinides, A.E. Texture segmentation based on higher order accumulators. Internal Report, Imperial College; 1991. Available from Prof. A.E. Constantinides, Electrical Engineering Department.
33. Brodatz, P. Textures – A Photographic Album for Artists and Designers. New York: Dover Publications; 1966.
34. SAS Institute Inc. SAS-BASE, SAS-STATISTICS, SAS-GRAPH. Cary, North Carolina 27511-8000, USA.