A. Materka, M. Strzelecki, R. Lerski, L. Schad: *Toward Automatic Feature Selection of Texture Test Objects for Magnetic Resonance Imaging*, RECPAD 2000, 11th Portugese Conference on Pattern Recognition, Porto, May 2000, 11-16.

Toward Automatic Feature Selection of Texture Test Objects for Magnetic Resonance Imaging

Andrzej Materka[1], Michał Strzelecki[1], Richard Lerski[2] and Lothar Schad[3]

[1]*Institute of Electronics, Technical University of Łódź*
Stefanowskiego 18
90-924 Łódź
POLAND
email: [materka][mstrzel]@ck-sg.p.lodz.pl

[2]*Medical Physics Department, Ninewells Hospital and Medical School*
Dundee DD1 9SY
UNITED KINGDOM
email: r.a.lerski@dundee.ac.uk

[3]*Deutsches Krebsforschungszentrum Abt. Radiologie*
Im Neuenheimer Feld 280
D-69120 Heidelberg,
GERMANY
email: l.schad@dkfz-heidelberg.de

Abstract: *An attempt is made to develop automatic techniques for the selection of image features for texture analysis of magnetic resonance test objects (phantoms). The test objects are designed for standardization of in vivo magnetic resonance imaging. They are made of reticulated foam embedded in agarose gel. Different porosity foam materials are used to manufacture the phantoms. The MR images of the foam phantoms are split into classes differing by the foam pore size. For any image texture, a large number of its statistical features can be computed. Since not all of them are carrying the information necessary for texture classification, there is a need to select those that have the largest discriminative power. Two feature-selection techniques are tested, based respectively on the value of Fisher coefficient F and combined probability of classification error and average correlation between features. Early experimental results are presented and discussed using a series of MR-scanned phantom images. The effect of MRI slice thickness on F is demonstrated as well.*

Keywords: *Feature selection, texture analysis, MRI*

1. INTRODUCTION

The traditional obstacles in obtaining quantitative description of the state of human body – invisible internal structures, inaccessible tissue, fast and complicated evolution in time and space, and traditional qualitative analysis methodology (careful but descriptive observation of individuals and population) – no longer seem impossible barriers [1]. Among the most significant technological advances that have contributed to the remarkable change from qualitative to quantitative medical diagnosis and therapy are imaging techniques. Especially magnetic resonance imaging (MRI) has made it possible to see the interior structures of the body and watch them without harm to the living subjects. There is a need for quantitative description of the MR images to stimulate more repeatable and objective medical diagnosis. One of the characteristics that are expected to carry on the diagnostic information is the image texture [9].

Studies have been carried out [8] to investigate whether texture measurements are transportable between magnetic resonance centers and to make firm conclusions as to the machine settings and sequence selection required. At present, development of quantitative methods of texture analysis of magnetic resonance images is the subject of COST B11 European Community project scheduled for the years 1998-2002 [7]. The aim of this project is to develop methods, which would allow reliable discrimination of different kinds of tissue in MR images, independent of scanner type and place of its installation.

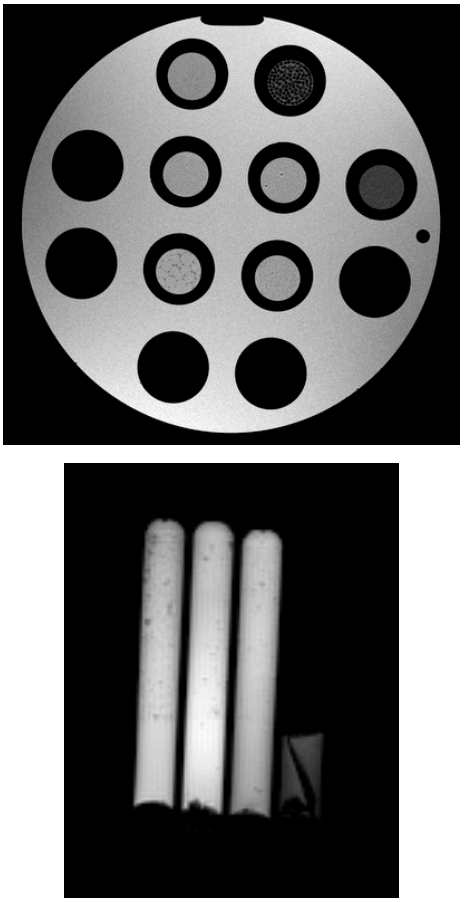


Figure 1. MR images of glass tubes filled with foam material in two cross-section views.

2. TEST OBJECT IMAGES

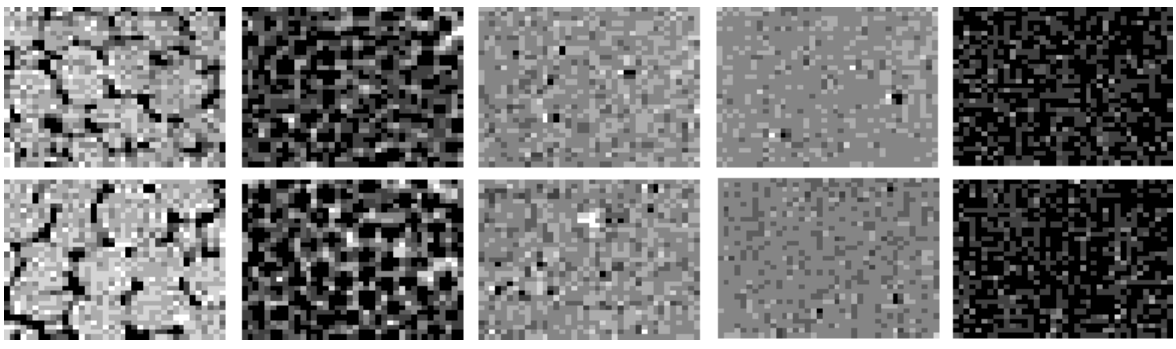
The use of texture analysis in magnetic resonance imaging requires the availability of texture test objects (phantoms) for use in standardization of *in vivo* measurements. Four physical phantoms were manufactured in Medical Physics Department, University of Dundee, Scotland. Three of them are in the form of glass tubes filled with different-porosity reticulated foam, one tube is filled with glass beads (Fig. 1). The foams and the space between glass beads are stuffed with agarose gel that possesses a relatively long value of magnetic resonance T2 response [8]. The tubes are sealed properly to prevent the water included in the gel from evaporation. A series of magnetic resonance images of the phantoms were recorded using a Siemens Magnetom 1.5-Tesla scanner at the German Research Cancer Center, Heidelberg, Germany. The images represent cross-sections of the foam- and glass beads-filled tubes, taken at field of view of 100 mm×100 mm, constant number of image pixels (256×256), different values of slice thickness (2 mm and 4 mm) – all acquired at 5 different positions along the tube axis. As a result, 4 different texture classes were obtained with five samples in each class, in this initial study. Each texture sample was an image containing 38×33 pixels. Example of the MR textured images is presented in Figure 2.

3. MATERIAL AND METHODS

A number of subroutines in Matlab and a specialized MS Windows application program MaZda [7] were written to compute a variety of texture features (parameters), including those computed from co-occurrence, run-length and gradient matrices [5]. The programs were applied to the recorded MR images to compute texture features and thus characterize texture properties.

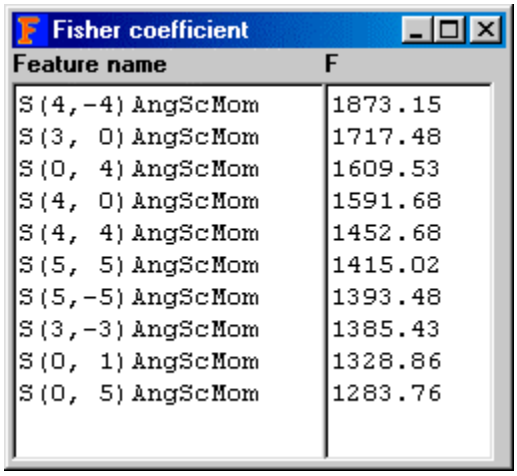
For each texture sample, the following 245 features were calculated:

- GR: 5 gradient-based features (absolute gradient mean, variance, skewness, kurtosis, and percentage of non-zero gradients),
- RL: 20 run-length matrix-based features (short run emphasis inverse moment, long run emphasis moment, gray level nonuniformity, run length nonuniformity and fraction of image in runs, separately for horizontal, vertical, 45° and 135° directions),
- CO: 220 co-occurrence matrix based features [11 features defined in [4] calculated for matrices constructed for five distances between image pixels (d=1, 2, 3, 4 and 5), and for the four directions as in the case of RL features].

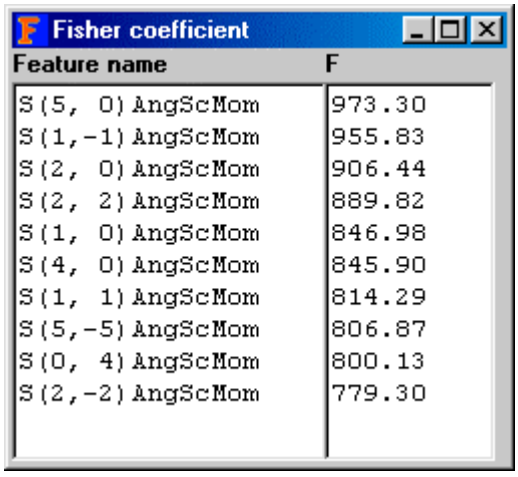


a1 a2 a3 a4 b

Figure 2. MR phantom images: a1 – foam, large pore size; a2 – glass bead; a3 – foam, medium pore size; a4 – foam, small pore size; b – background noise. Slice thickness: upper row – 4 mm, lower row – 2 mm.



a)



b)

Figure 3. Tables of features selected for 5 texture classes according to the value of F coefficient. MRI slice thickness: a) 2mm, b) 4mm.

Prior to feature calculation, mean value μ and standard deviation σ of each image was computed. These values were then used to normalize images – by quantizing their intensity in the range from μ-3σ to μ+3σ [11] to 64 gray levels (6-bit word-length).

To quantitatively evaluate the features ability to separate different texture classes, two different techniques have been used in this study. One of them is based on Fisher F coefficient [13], which is the ratio of mean-squared between-class distance D^2 (computed between the class means $\boldsymbol{\mu}_k$, k=1,2,...,K) to the mean of mean-squared within-class distances V_k^2 (computed between the samples of class k and the corresponding class mean $\boldsymbol{\mu}_k$)

$$F = \frac{D^2}{V^2} = \frac{\dfrac{1}{1-\sum_{k=1}^{K} P_k^2} \sum_{k=1}^{K}\sum_{j=1}^{K} P_k P_j \left|\boldsymbol{\mu}_k - \boldsymbol{\mu}_j\right|^2}{\sum_{k=1}^{K} P_k V_k^2} \quad (1)$$

where P_k denotes the *a priori* probability of data belonging to texture class k. The other one uses the combined measures of classification error (POE – probability of error) and average correlation coefficient (ACC) between features so far selected [12]. A package of MS Windows 9x/NT programs was written to implement the two automatic methods for feature subset selection.

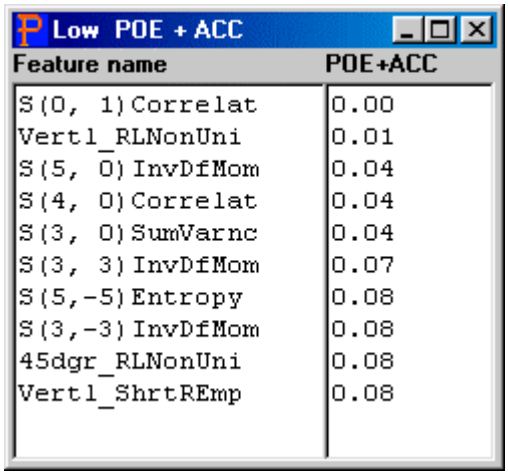
For feature vector classification, the nearest neighbor, k-NN, technique [2] was used. Due to the small size of the data (25 sample vectors, 5 samples per class), only one neighbor (1-NN) was considered. Finally, small size artificial neural networks (ANN) were trained on the available data to make the data nonlinear transformation and to classify the textures. The ANNs considered were single hidden layer feedforward networks with sigmoidal processing elements [6].

4. RESULTS AND DISCUSSION

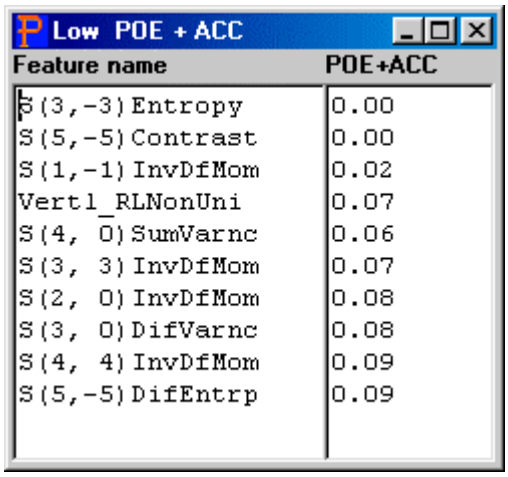
Figure 3 demonstrates a list of first ten texture features selected by the program according to descending Fisher coefficient value computed from (1) with K = 5. For further reference, the features chosen for slice thickness of 2 mm (Fig. 3a) are denoted as feature set 1, FS1, and those chosen for slice thickness of 4 mm (Fig. 3b), are designated as FS2. Interestingly, for both MRI slice thickness values, the only feature selected to form FS1 and FS2 was the so-called angular second moment, AngScMom, [4],[5], derived from the co-occurrence matrix. In each feature case, the CO matrix chosen was computed for different separation and orientation of pixel pairs. [As an example, S(4,-4) denotes every pair of pixels that are separated by 4 sampling intervals in the horizontal direction, and –4 intervals in the vertical direction. In other words, these particular pixels are located on a diagonal line, oriented at 45 degrees from horizontal direction.] The inter-pixel distances selected by the program cover the whole range of distances available in the program, from 1 to 5. This is in agreement with the fact that texture classes under investigation cover a broad range of texture primitive element sizes (corresponding roughly to the foam pore size).

Two lists of ten texture features chosen by the POE+ACC technique are shown in Fig. 4, respectively for MRI slice thickness of 2 mm and 4 mm. Consequently, they are denoted by FS3 (Fig. 4a) and FS4 (Fig. 4b). Most of these features have their origin in the CO matrix of different pixel-pair distance and orientation. However, the run-length and absolute gradient matrices are also represented.

Only two features appeared simultaneously in FS3 and FS4 feature sets. They are "Vertical Run-Length Nonuniformity" (Vertl_RLNonUni) and (3,3) CO matrix-derived "Inverse Differential Moment" (InvDfMom) for the POE+ACC technique. In the case of Fisher coefficient-based selection (FS1 and FS2), the CO-matrix feature "Angular Second Moment" (AngScMom) repeated itself, obtained for two



a)



b)

Figure 4. Tables of features selected for 5 texture classes according to the minimum weighted sum of classification error and mean inter-feature correlation coefficient (POE+ACC method). MRI slice thickness: a) 2mm, b) 4mm.

different pixel separation vectors (0,4) and (5,-5). The set of these 4 features was also used, as the basis for fifth data set for texture classification, FS5, for 2-mm slices and set FS6, for 4-mm slices.

First, an attempt was made to classify the data vectors corresponding to feature sets FS1 – FS6. The results are presented in Table 1. One can see that the numbers of missclassified raw (unprocessed) data vectors are significant.

Table 1. Results of 1-NN classification of raw data vectors

Data set	Number of errors
FS1 (2 mm, Fisher)	3
FS2 (2 mm, POE+ACC)	8
FS3 (4 mm, Fisher)	3
FS4 (4 mm, POE+ACC)	8
FS5 (2 mm)	11
FS6 (4 mm)	10

To check whether optimal linear data transformation can help to discriminate the texture classes under consideration, the linear discriminant analysis (LDA) was applied to the collected data vectors [3]. These vectors were projected on the eigenvectors of the $\mathbf{C}_W^{-1}\mathbf{C}_T$ matrix, where $\mathbf{C}_W$ is the within-class scatter matrix and $\mathbf{C}_T$ is the total scatter matrix. In the case of 10-dimensional features from sets FS1-FS4, the features obtained after projection were 4-dimensional, which is in agreement with theory (4 = 5-1, where 5 is the number of classes in the considered case). The LDA features were used to form new data sets, which were then classified using the 1-NN technique. The results of the classification are presented in Table 2. One can notice a significant reduction in the classification error. This means that features selected to describe the texture under consideration carry the information necessary to discriminate the 5 texture classes. Moreover the 5 classes seem to be separable; however, more experiments are needed to confirm this, based on larger data sets.

Table 2. Results of 1-NN classification of LDA vectors

Data set	Number of errors
FS1 (2 mm, Fisher)	2
FS2 (2 mm, POE+ACC)	0
FS3 (4 mm, Fisher)	0
FS4 (4 mm, POE+ACC)	0
FS5 (2 mm)	0
FS6 (4 mm)	2

To see whether other classification techniques would confirm the above findings, one more experiment was conducted using ANNs with a small number of weights. To further reduce the number of weights and thus avoid the overtraining problem, the 4-dimensional LDA features were used as ANN inputs, instead of 10-dimensional raw data vectors. The ANN was used to project the input to a 2-dimensional space, called the nonlinear discriminant analysis (NDA) space [10]. As an effect, further data reduction was possible. Figure 5 shows the obtained texture class clusters on the NDA plane. Apparently, all the 5 clusters are linearly separable, giving the evidence that the MRI phantoms can be firmly distinguished based on their texture.

It is interesting to note that Fisher coefficient computed for 2 mm MRI slices is about two times larger compared to its value for 4 mm images. This is illustrated in Table 3. This reflects the smoothing effect of small structural details of the foam material as they are averaged over a larger volume of 4-mm thick MRI slices compared to thinner, 2-mm slices. In other words, the discrimination measure is higher in

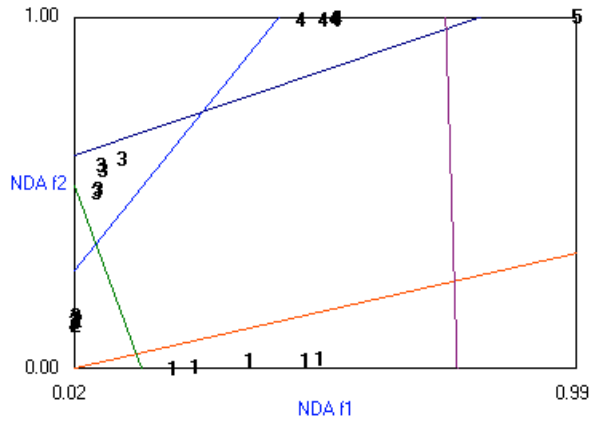


Figure 5. Nonlinearly projected features on a 2-dimensional NDA feature space. Segments of straight lines separate the five texture classes in the new feature space.

the case of 2-mm slice images, because for thinner slice, pores located in deeper layers distort the original (cross-sectional) texture structure to a lower extent. The effect of this distortion is especially visible in Fig. 2, where images from upper (4 mm) and lower (2 mm) rows can compare to each other. The experiments conducted show that texture analysis is a useful tool to evaluate the influence of MRI scanner parameters on diagnostic contents of MR images.

Another observation drawn from Table 3 is that the value of F is larger for FS1 and FS3 as compared to FS2 and FS4. This can be explained by the fact that it is the F coefficient value which is maximized by the choice of features in the first feature selection technique described in Section 2. The POE+ACC method does not take the F value into account directly, so the resulting values of (1) are generally smaller for the features selected by this technique. There is no evidence available by now to demonstrate superiority of any of these techniques to each other – further analysis is needed based on larger data sets. One should remember that the sample size is very small in this study and further justification of the possibility is necessary, based on more extensive experimental material.

Table 3. Fisher coefficient for raw data sets

Data set	Fisher coefficient
FS1 (2 mm, Fisher)	1485.5
FS2 (2 mm, POE+ACC)	84.6
FS3 (4 mm, Fisher)	857.0
FS4 (4 mm, POE+ACC)	28.0

5. CONCLUSIONS

An attempt has been made to evaluate the effectiveness of statistical parameters as texture features to discriminate between different test objects for magnetic resonance imaging. Since a vast number of different feature definitions has been developed and described in the literature, there is a need for developing techniques for best feature selection – for the task of texture classification. Two methods of automatic feature selection were considered in the paper and tested – based on the value of Fisher coefficient and based on minimum classification error and minimum correlation between selected features. Both techniques are effective in terms of generating subsets of features that can be used to classify the textures with a very small error. Linear and nonlinear discriminant analyses are useful to form compact data sets for classification.

The experiments with a set of measured images of phantom objects showed that classification of physical objects of different internal structure is possible based on MRI texture parameters. The experiments conducted with MR images of different slice thickness indicate that texture analysis is a useful tool to evaluate the influence of MRI scanner parameters on diagnostic contents of MR images. This can be of much value to quantitative description of the MR imaging process and can contribute to more repeatable and objective medical diagnosis.

For the future, the following investigations are planned:

- consideration of new texture features (e.g. wavelet and mathematical morphology based features),
- comprehensive analysis of noise, slice thickness and field of view MR parameter influence on classification accuracy and selection of features that would be weakly dependent on noise,
- extending the results to texture classification of biological tissue,
- analysis of efficiency of other known techniques for feature selection, data preprocessing and classification applied to the task of MRI interpretation.

Acknowledgment: This work was performed within the framework of COST B11 European project. It was supported in part by British-Polish Joint Research Programme.

REFERENCES

[1] R. Barr, "Biomedical engineers", *IEEE Spectrum*, January 1999, pp. 82-83.

[2] R. Duda, P. Hart, *Pattern Classification and Scene Analysis*, Wiley 1973.

[3] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, 1990.

[4] R. Haralick, K. Shanmugam and I. Dinstein, "Textural Features for Image Classification", *IEEE Trans. Systems Man Cybernetics*, **3**, 6, 1973, pp. 610-621.

[5] R. Haralick, Statistical and Structural Approaches to Texture, *Proceedings IEEE*, **67**, 5, 1979, 786-804.
[6] Hecht-Nielsen R. *Neurocomputing*, Addison-Wesley 1990.
[7] Internet 1999, http://phase.pki.uib.no/~costb11/
[8] R. Lerski, L. Schad, "The Use of Reticulated Foam in Texture Test Objects for Magnetic Resonance Imaging", *Magnetic Resonance Imaging* **16**, 9, 1998, pp. 1139-1144.
[9] R. Lerski, K. Straughan, L. Schad, D. Boyce, S. Bluml, I. Zuna, "MR Image Texture Analysis – An Approach to Tissue Characterization", *Magnetic Resonance Imaging* **11,** 1993, pp. 873-887.
[10] J. Mao, A. Jain, "Artificial Neural Networks for Feature Extraction and Multivariate Data Projection", *IEEE Trans. on Neural Networks*, 6, 2, 1995, pp. 296-316.
[11] A Materka, M Strzelecki, R Lerski, L Schad, "Evaluation of Texture Features of Test Objects for Magnetic Resonance Imaging", *Infotech Oulu Workshop on Texture Analysis in Machine Vision*, Oulu, June 1999, Finland, pp. 13-19.
[12] A. Mucciardi, E. Gose, "A Comparison of Seven Techniques for Choosing Subsets of Pattern Recognition Properties", *IEEE Transactions on Computers*, 20, 9, 1971, pp. 1023-1031.
[13] J. Shürmann, *Pattern Classification*, John Wiley and Sons, New York, 1996.