A. Materka, M. Strzelecki, R. Lerski, L. Schad: *Evaluation of Texture Features of Test Objects for Magnetic Resonance Imaging*, June 1999, Infotech Oulu Workshop on Texture Analysis in Machine Vision, Oulu, Finland, 13-19.

FEATURE EVALUATION OF TEXTURE TEST OBJECTS FOR MAGNETIC RESONANCE IMAGING

Andrzej Materka, Michał Strzelecki
Institute of Electronics, Technical University of Łódź
Stefanowskiego 18, 90-924 Łódź, Poland.
[materka][mstrzel]@ck-sg.p.lodz.pl
Richard Lerski
Medical Physics Department, Ninewells Hospital and Medical School
Dundee DD1 9SY, United Kingdom
r.a.lerski@dundee.ac.uk
Lothar Schad
Deutsches Krebsforschungszentrum Abt. Radiologie
Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany
l.schad@dkfz-heidelberg.de

Abstract. Texture analysis of test object (phantom) images for standardization of *in vivo* magnetic resonance imaging is considered in this paper. The test objects are made of reticulated foam embedded in agarose gel. Different porosity foam materials are used to manufacture the phantoms. Both optical and MR images are analyzed, split into classes differing by the foam pore size. To characterize the image texture, a number of its first- and second-order statistical features is computed. The usefulness of the features to object class discrimination is evaluated using the ratio F of between-classes variance to within-classes variance, and multidimensional analysis of variance. The effect of noise and MR slice thickness on F is investigated.

1. Introduction

Studies have been carried out (Lerski 1993) 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. Development of quantitative methods of texture analysis of magnetic resonance images is now the subject of COST B11 European Community project scheduled for the years 1998-2002 (Internet 1999). 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. There is an expectation that the texture analysis technique will contribute to more objective and repeatable medical diagnosis.

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. They are in the form of glass tubes filled with different-porosity reticulated foam. The foam is stuffed with agarose gel that possesses a relatively long value of magnetic resonance T2 response (Lerski 1998). The tubes were 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-filled tubes, taken at different field of view (100 mm×100 mm and 200 mm×200 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. Example of the MR textured images is presented in Figure 1. Phantom images are analyzed in the current stage of investigation, before *in vivo* experiments on tissue texture analysis – planned for the future – are carried on.

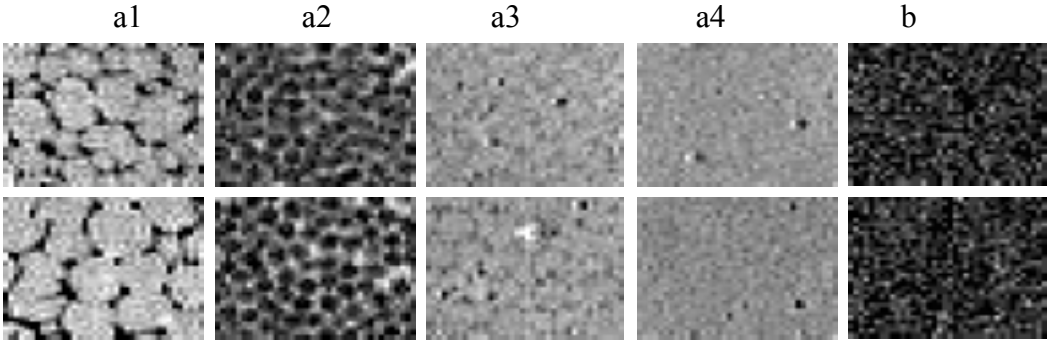


Figure 1 MR phantom images: a1 – foam, large pore size; a2 – glass bead; a3 – foam, medium pore size; a4 – foam, small pore size; b – background noise.
Slice thickness: upper row – 4 mm, lower row – 2 mm.

Independently, optical images of the reticulated foam materials were digitally recorded, for comparison with MR images. They contain scans of cross-section of two different-porosity foams. Obtained 8-bit images of size 175x175 pixels are shown in Figure 2. From each optical image, 42 non-overlapping samples of size 23x23 pixels were taken, resulting in 2 texture classes, each of 42 samples.

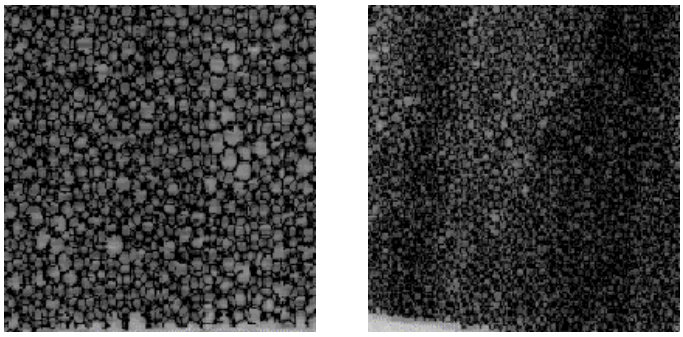


Figure 2 Optical images of reticulated foam materials:
Foam1 – large pore size, Foam2 – medium pore size.

3. Results and discussion

A number of subroutines in Matlab and a specialized MS Windows application program MaZda (Internet 1999) were written to compute a variety of texture features (parameters), including first-order (histogram-based), second-order (computed from co-occurrence and run-length matrices) (Haralick 1979), gradient, and autoregressive (AR) model-based features (Hu 1994). The programs were applied to the recorded MR and optical images to compute texture features and thus characterize texture properties. In the case of MR images, the effect of noise was taken into account by adding Gaussian noise of specified standard deviation to image samples and then computing the features. To investigate any feature ability to discriminate between different pore-size textures, the following *F* coefficient was used:

$$F = \frac{D}{V} \qquad (1)$$

that represents the ratio of between-classes feature variance *D* to within-classes feature variance *V* (Shürmann 1996).

For each sample (region of interest – ROI) of an optical image, the following 254 features were calculated:

- H: 9 histogram-based (mean, variance, skewness, kurtosis and five histogram percentiles for 1%, 10%, 50%, 90%, and 99%: #1 – #9),
- GR: 5 gradient-based features (absolute gradient mean, variance, skewness, kurtosis, and percentage of non-zero gradients: #10 – #14),
- 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: #15 – #34),
- CO: 220 co-occurrence matrix based features (11 features defined in (Haralick 1973) 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: #35 – #254.

Except for the histogram-based features, each ROI image was quantized to 64 gray levels (6-bit word-length) prior to computation of the texture parameters.

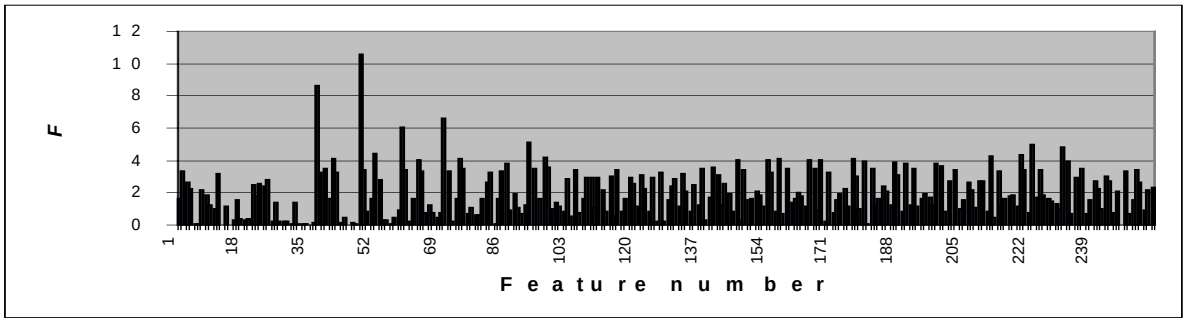


Figure 3 *F* coefficient for H, GR, RL and CO features (no ROI normalization).

For each of the above-mentioned features, *F* coefficient was computed to express the possibility to separate the two foam classes based on a given feature. As presented in Figure 3, for raw images, i.e. with no image normalization within ROIs, only 4 features from the whole set (#37, #48, #59, and #70) represent relatively high value of *F* coefficient (e.g. $F \geq 6.0$). They are correlation coefficients calculated for the co-occurrence matrix determined at $d = 1$, for the four main directions. Other features possess lower *F* values, which means that they are not very useful to make distinction between the two classes of the foam texture, cf. Table 1.

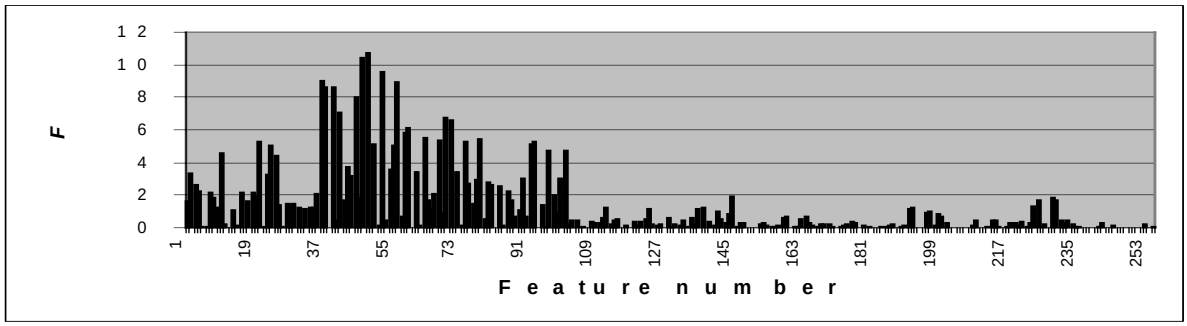


Figure 4 *F* coefficient for H, GR, RL and CO features ('±3 sigma' ROI normalization).

To investigate whether image normalization has an influence on feature ability to allow discrimination between the image classes, two normalization schemes were considered. For both schemes, the image histogram was first computed within each ROI. Then, the image mean μ and standard deviation σ were found. For the '±3 sigma' scheme, the image intensity

levels were limited to the range from a minimum of $f_{min}=\mu-3\sigma$ to a maximum of $f_{max}=\mu+3\sigma$. The intensity range (f_{max}–f_{min}) was then quantized using 6-bit word-length prior to computation of GR, RL and CO parameters. On the other hand, for the '1% – 99%' scheme, the values of f_{min} and f_{max} were determined as corresponding to, respectively, 1% and 99% of cumulative image histogram within ROI.

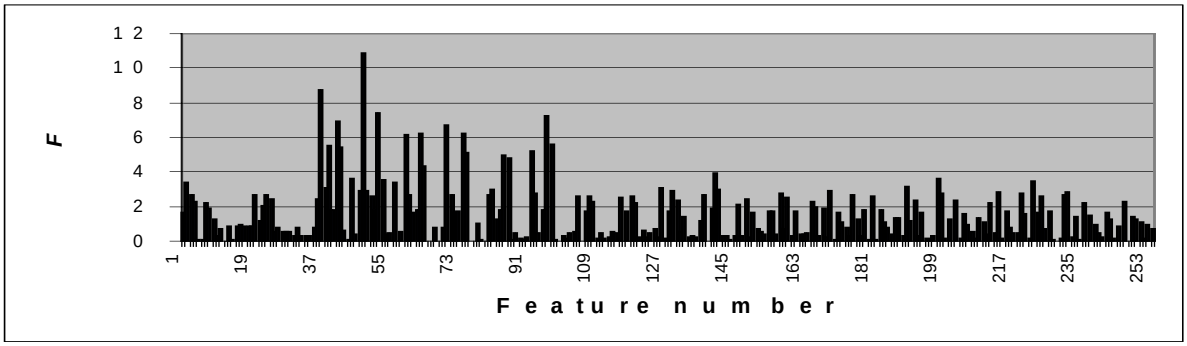


Figure 5 *F* coefficient for H, GR, RL and CO features ('1%–99%' ROI normalization).

The *F* coefficient distribution among the different texture features, obtained for the '±3 sigma' and '1%–99%' normalization schemes is illustrated in Figures 4 and 5, respectively. Indeed, for '±3 sigma' scheme, the number of features that have high *F* ($F \geq 6.0$) increased significantly to 12. An intermediate number of 9 such features was obtained for '1%–99%' scheme. Numerical results of this experiment are presented in Table 1.

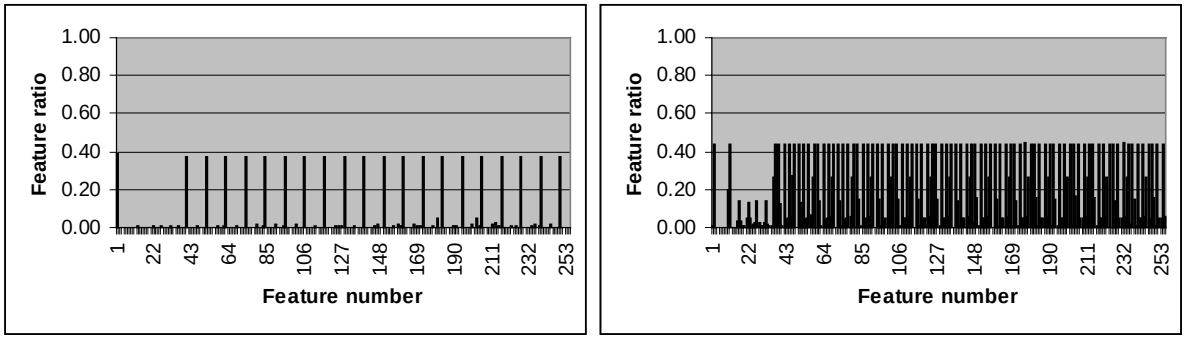
Table 1
F coefficient for different normalization schemes (shaded areas: $F>6$).

No.	Feature definition	Feature number	*F*, no normalization	*F* '±3 sigma'	*F* '1%–99%'
1	(1,0) Contrast	36	0.2	9.0	2.4
2	(1,0) Correlation	37	8.7	8.6	8.8
3	(1,0) Inverse Differential Moment	39	3.5	8.7	5.5
4	(1,0) Sum Variance	41	4.1	7.1	7.0
5	(1,0) Differential Entropy	45	0.0	8.0	3.6
6	(0,1) Contrast	47	0.1	10.4	2.9
7	(0,1) Correlation	48	10.6	10.7	10.9
8	(0,1) Sum Variance	52	4.4	9.6	7.5
9	(0,1) Differential Entropy	56	0.1	9.0	3.4
10	(1,1) Correlation	59	6.1	6.2	6.1
11	(1,1) Sum Variance	63	4.0	5.6	6.3
12	(1,-1) Contrast	69	0.8	6.7	0.8
13	(1,-1) Correlation	70	6.6	6.6	6.7
14	(1,-1) Sum Variance	74	4.1	5.3	6.3
15	(0,2) Sum Variance	96	4.2	4.7	7.3

It is evident that the number of useful features depends significantly on image normalization. To explain this effect, one should refer to image properties as seen in Figure 2. Namely, the images investigated, especially 'Foam2', show some nonuniformity of their local mean and variance. It can be found that relative standard deviation σ_μ of image mean μ, computed over 48 ROIs, is equal to $\sigma_\mu/\mu=12.2\%$ for 'Foam1' and as much as $\sigma_\mu/\mu=24.2\%$ for 'Foam2'. Similarly, the corresponding ratios related to image variance are equal to 15.8% for 'Foam1' and 29.1% for 'Foam2'. At the same time, *F* coefficient for image mean μ is equal to 1.6 and that for image variance σ^2 equals to 3.4. One can then expect that if there exist texture features, which possess high correlation to μ and σ^2, and image is not normalized, then such features will demonstrate non-zero values of *F* even if they do not carry any information

about texture properties other than μ and σ^2. Such features will be redundant in a given application. Moreover, high correlation to μ and σ^2 may mask a feature ability to discriminate the texture classes.

To find out whether there are indeed features highly correlated to μ and σ^2 (within the feature set under consideration) two numerical experiments were carried out – one with modified image mean, and the other with a modified variance. For each texture feature, the ratio of its value after mean (variance) modification to the value before modification was calculated, as presented in Figure 6.



(a) (b)

Figure 6 The effect of image mean (a) and variance (b) on texture features.

The increase in image mean corresponding to Figure 6a is equal to 0.39 of its non-modified value. Similar increase (0.37) is observed for the CO Sum Average features, regardless of the distance *d* and direction used for CO matrix computation. These features are represented by 20 respective bars in Figure 6a, numbered as 40, 51, ..., 249. The GR (#10-#14) and RL (#15-#34) features did not depend on μ, as can be expected. The histogram percentile features (#5 - #9) were set to zero to neglect the effect of slight intensity clipping due to the modifications.

Figure 6b indicates that many more features depend on σ^2 (#2) compared to μ. The increase in σ^2 shown in this Figure equals to 0.44. The same increase is observed for gradient variance (#11). The gradient mean is increased by a factor of 0.2 (#10). As far as RL features are concerned, the increase of 0.14 is measured for the so-called Run Length Nonuniformity (Haralick 1973), regardless of the direction (#16, #21, #26, and #31). Most of the CO-derived features demonstrate dependence on σ^2. The same increase of 0.44 is obtained for CO Contrast (#36), CO Sum of Squares (#38), CO Sum Variance (#41) and CO Difference Variance (#44). The CO Angular Second Moment (#35) increased by 0.27, and the CO Inverse Difference Moment (#39) changed by a factor of 0.12. Only CO Correlation (#37) is independent of image variance. The CO features numbered from #35 to #45 correspond to all pairs of image points that are a vector (1,0) apart from each other. The whole pattern of CO feature dependence on σ^2 repeats in Figure 6b with a period of 11, until the features obtained for the displacement vector (5,-5) that are numbered from #244 to #254.

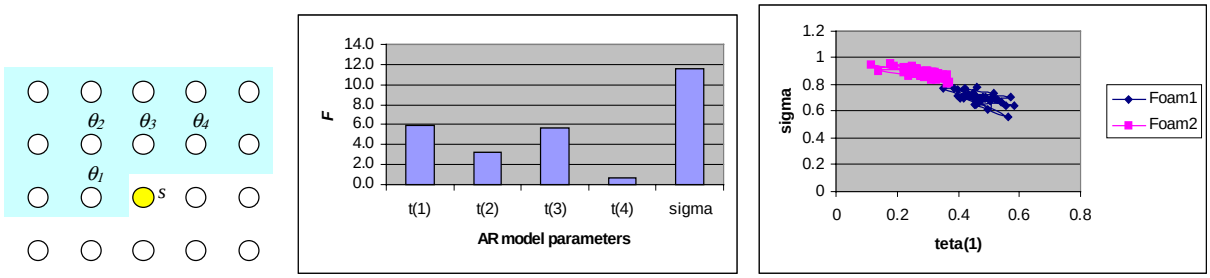
More detailed analysis shows that standard deviation of the mean of 'Foam1' equals to 7.8, 0.2, and 4.5, respectively for 'no normalization', '±3 sigma', and '1%-99%' schemes. The corresponding figures for 'Foam2' are equal to 11.3, 0.2, and 3.5. Thus '±3 sigma' scheme provides the best stabilization of the image mean value within the (f_{max}–f_{min}) window. This results in the highest number of the discriminative features (Figure 4), thanks to elimination of the effect of their correlation to mean and variance [μ=0 and σ^2 is constant relative the

(f_{max}–f_{min}) intensity range for the '±3 sigma' normalization scheme]. Table 1 indicates that CO Correlation does not indeed depend on image normalization.

To extend the set of texture features beyond those discussed so far, a small-neighborhood AR model parameters θ_1, θ_2, θ_3, θ_4, and σ_{ar} were used (Hu 1994)

$$f_s = \sum_{r \in N_s} \theta_r f_r + e_s \qquad (2)$$

where f_s is image intensity at site s, e_s denotes an i.i.d. driving noise of standard deviation σ_{ar}, θ_i, $i=1,...,4$ represent selected pixel-to-pixel relationship, σ_{ar} is the noise standard deviation and N_s is a neighborhood of s. The image sample mean and variance were both normalized to, respectively, 0 and 1 in the case of AR model identification. The F values obtained for AR parameters are presented in Figure 7b. In this case, three AR parameters have relatively high F values: θ_1, θ_3, and σ_{ar}. They can be used for foam textures discrimination. This is diagrammed in Figure 7c. As can be observed, the σ_{ar} feature alone is sufficient to discriminate the two optical textures.



(a) (b) (c)

Figure 7 AR model neighborhood (a), F coefficient for AR model parameters (b), scatter-plot of two AR model features (c) – optical foam images.

In this preliminary study, the number of MR image samples for each of the 5 texture classes illustrated in Figure 1 is limited to 5 only (25 images for each slice thickness were available, total of 50 images in the whole experiment). This eliminates the possibility of training e.g. a neural-network classifier, without its overtraining. Such attempts have then been postponed for future study, when more MR image samples are collected. Nevertheless, with the material at hand, it was still possible to draw some consistent conclusions with regard to the effects of slice thickness and noise on the class separability measure. The above-defined 5-parameter AR model was used to characterize the measured MR images of the test objects.

Figure 8 shows the F coefficient as a function of noise standard deviation (added to the images shown in Figure 1) – calculated for all five AR model parameters of MR images. As can be observed, discrimination ability decreases with increasing noise standard deviation. Also, the discrimination measure is higher in 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 Figure 1, where images from upper and lower image rows can compare to each other.

It follows from the numerical experiments that it is possible to classify the 5 MR texture classes with no error, based on the small-neighborhood AR model parameters. However, 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. The discrimination measure, as illustrated in Figure 8, seems to show monotonous deterioration with both slice thickness and noise. Both of these parameters of the magnetic resonance

imaging process have the effect on the time of MR image measurement, and thus its cost and independence of patient body movement. It is expected that some guidelines can be formulated in future as to the time required to ensure image quality that would guarantee texture classification with a prescribed permissible error.

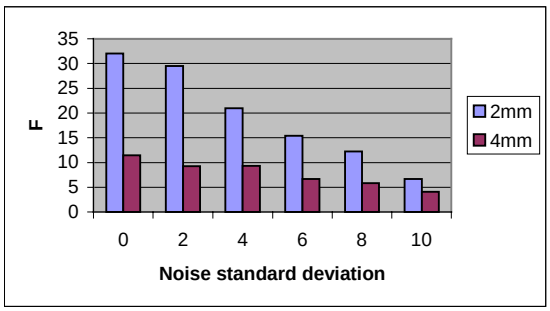


Figure 8 *F* coefficient as a function of additive noise standard deviation for AR model parameters (MR images).

4. 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. Surprisingly, only a few among more than 200 popular features are useful to distinguish the otherwise quite distinct (at least for humans) textures. This indicates the need for carrying on research work on better understanding of texture properties and for finding new feature definitions that would provide means for firm discrimination of different images of biological origin. The significance of image normalization prior to texture parameter computation has been demonstrated. This early study shows usefulness of AR model parameters for both MR and optical texture discrimination. 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 influence on classification accuracy and selection of features that would be weakly dependent on noise,
- further development of MaZda software used for feature calculation,
- development of feature selection methods for MR texture test objects,
- extending the results to texture classification of biological tissue.

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

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

Internet 1999, http://phase.pki.uib.no/~costb11/

Y. Hu and T. Dennis, "Textured Image Segmentation by Context Enhanced Clustering", *IEE Proc.-Visual Image and Signal Processing*, **141**, 6 (1994) 413-421.

R. Lerski, L. Schad, "The Use of Reticulated Foam in Texture Test Objects for Magnetic Resonance Imaging", *Magnetic Resonance Imaging* **16**, 9 (1998) 1139-1144.

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) 873-887.

J. Shürmann, *Pattern Classification*, John Wiley and Sons, New York, 1996.