Medical Physics
The International Journal of Medical Physics Research and Practice

Phantoms for texture analysis of MR images. Long-term and multi-center study

Daniel Jirák, Monika Dezortová, and Milan Hájek

Citation: Medical Physics **31**, 616 (2004); doi: 10.1118/1.1646231
View online: http://dx.doi.org/10.1118/1.1646231
View Table of Contents: http://scitation.aip.org/content/aapm/journal/medphys/31/3?ver=pdfcov
Published by the American Association of Physicists in Medicine

Phantoms for texture analysis of MR images. Long-term and multi-center study

Daniel Jirák,[a] Monika Dezortová, and Milan Hájek
Department of Diagnostic and Interventional Radiology, Institute for Clinical and Experimental Medicine, Videnska 1958/9, Prague, Czech Republic 140 21, Czech Republic

(Received 18 April 2003; revised 17 December 2003; accepted for publication 17 December 2003; published 26 February 2004)

The application of texture analysis (TA) in magnetic resonance imaging (MRI) requires the availability of texture phantoms for use in the standardization of *in vivo* measurements. The aims of our study were (a) to develop a new type of phantoms suitable for MRI and TA and test their long-term stability; (b) to optimize the choice of texture parameters describing the phantoms; (c) to compare different MR imagers according to texture parameters in a multi-center study. A long-term study performed at 4.7 T proved that the developed phantom based on polystyrene spheres and an agar gel solution is stable at least 12 months. This phantom, with nodular patterns, was found useful for the modeling of structural differences. The comparison of TA parameters at 4.7 and 7 T proved that the same parameters can be used for the separation of structures. The proposed algorithm of the selection of TA parameters shows that there exists a part of texture parameters which can be measured with high reproducibility (1–3%); on the other hand, their absolute values can differ by more than 30% if the textures differ. Results obtained from the multi-center study of whole body MR imagers show the wide variation in the misclassification rates at the different sites and point out the importance of the set up of MR sequences. © *2004 American Association of Physicists in Medicine.* [DOI: 10.1118/1.1646231]

Key words: Long-term stability, magnetic resonance imaging, multi-center study, phantom design, texture analysis

I. INTRODUCTION

Tissue in digital images (represented in gray scale) is characterized by texture as a distribution of signal intensity represented by a pattern of brightness and darkness. There are several mathematical and statistical methods that can eliminate possible subjective errors in the classification of different tissues and help in the description of digital images. One of these methods is known as texture analysis (TA). Texture analysis is a method that uses mathematical parameters for describing images (optics, ultrasound, CT, MR images, etc.), so TA is able to describe images quantitatively.

MR images are available in digital form which is represented by a matrix of pixels. Their intensities are input for the calculation of TA parameters. Methods of TA are divided into two major categories: statistical and syntactic. Medical applications generally prefer the statistical methods of TA to find the interrelationships among pixels, and these procedures have been applied in different studies (brain, liver, etc.).[1–3] In this paper we deal with this approach.

There are two important points which are discussed within the TA evaluation process. The first one is concerned with the information content of texture parameters. The second one is concerned with the choice of parameters which accurately distinguish the texture of different tissues. In real situations both problems have to be solved together. Two different approaches can be used. The first searches for the parameters with the most discriminative power which are able to distinguish different textures in a set of images. Such procedures usually use feature reduction methods based on choosing parameters according to given mathematical criterion such as the Fisher technique (F-coefficient)[4] or the combined probability of classification error and average correlation coefficients among parameters (POE).[5]

The second approach supposes that texture parameters exist which can be measured with very high reproducibility and which yield significantly unique results when applied to different textures.

Mathematical analysis offers several hundred features for the description of images;[6–8] nevertheless, in practice we are able to use only a few of these. Therefore, it is very important to find them and to understand how they can be used for the correct classification of texture. As in many other branches of MR imaging, phantoms are needed for the evaluation of the possibilities of TA in MR imaging.

Comparisons of various MR imagers have been performed, but generally on anatomic tissue.[9,10] TA of phantoms can compare the MR imagers quantitatively. Phantoms are principally used for quality control of MR imagers.[11–13] The construction of phantoms for quality control usually represents a set of different objects and the comparison of real distances or shapes is compared with the results of MR imaging. Phantoms based on foam or tubes which were filled with a water solution with a contrast agent were developed for texture analysis,[14] but there were some problems with air bubbles in the phantoms (M. Hájek, D. Jirák, V. Herynek, ''Phantoms PSAG,'' presented as a personal communication

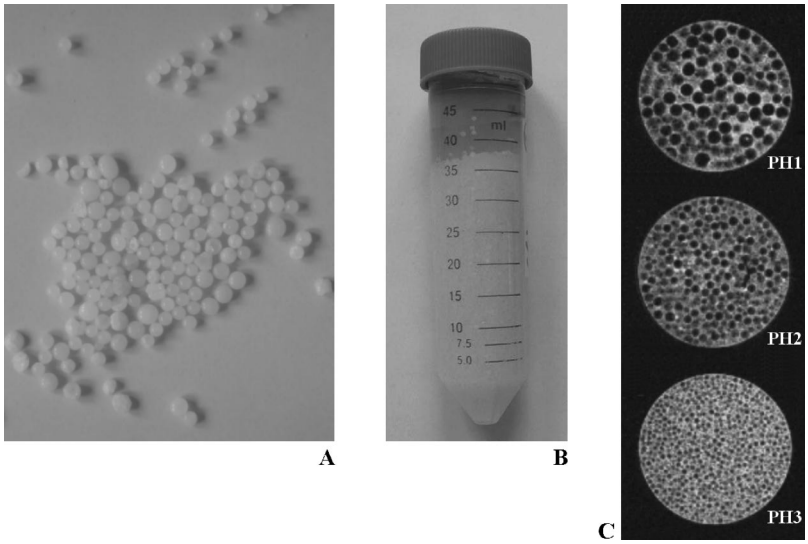

FIG. 1. PSAG phantom. Polystyrene spheres (obtained from the standard technological polystyrene production process) (a) are mixed with an agar solution and poured into a plastic tube (b). Different textures can be measured as is seen in MR images (PH1, PH2, and PH3) obtained from Bruker 4.7 T (c).

at COST B11. Quantitation of Magnetic Resonance Image Texture, Rennes, France, 2000). Therefore, we developed and tested a new type of phantom for texture analysis of MR images based on polystyrene spheres and an agar gel solution.

The aims of this study were as follows: (1) to develop new types of phantoms suitable for a texture analysis of MR images and test their long-term stability; (2) the testing of various techniques of an application of phantoms for the selection of TA parameters describing the various phantoms; (3) to check the possibility of separation of different structures (at the resolution limits) using TA which are measured by standard sequence with the standard parameters used in a clinical examination.

II. MATERIAL AND METHODS

A. PSAG phantoms

The PSAG phantoms with the nodular pattern were prepared using polystyrene spheres (PS) and an agar solution (AG).[15] Polystyrene spheres were obtained from the standard technological polystyrene production process (Kaučuk a.s., Czech Republic). Three groups of spheres with the following diameters were used for phantom construction: PH1 2.0–3.15 mm, PH2 1.25–2.0 mm, and PH3 0.8–1.25 mm. Polyethylene tubes of diameter 28 mm were filled in with spheres and mixed with a hot 4% water solution of agar (Fig. 1). 1 ml of 0.1% NaN3 (Merck & Co., Inc., USA) per 1 liter of agar (Penta, Czech Republic) was applied for microbiological stability. Phantoms were kept refrigerated ($7\pm2\,^{\circ}$C) and thermo-statted to the laboratory temperature before the experiments.

B. MRI protocols

MR tests of long-term stability of phantoms were performed at 4.7 T with a Bruker Biospec spectrometer equipped with a standard resonator coil (receiver and transmitter, inner diameter = 7 cm). Phantoms with two sphere diameters (PH1 and PH2) were measured over twelve months (once per month). A turbo-spin echo sequence RARE (Rapid Acquisition with Relaxation Enhancement) was used with these parameters: repetition time (TR) = 3000 ms, echo time (TE) = 62.7 ms, rare factor = 8, matrix (MTX) = 256×256, field of view (FOV) = 3.5 cm, number of axial slices = 13, slice thickness = 1 mm, slice distance = 5 mm, number of experiments (NEX) = 4, resolution = 0.137×0.137 mm per pixel, laboratory temperature = 18 °C.

The experiments at 7 T were done with two phantoms (PH1 and PH2) with a SMIS experimental imager (University of Antwerp) equipped with a resonator coil. A turbo-spin echo sequence RARE was used as a standard sequence with the same parameters as at the 4.7 T experiments. The sequence for 7 T experiments differed only in the TE parameter (TE = 30 ms) because shorter TE for the same sequence as at 4.7 T was not available.

Five whole body clinical imagers were used for measurement of three phantoms (PH1, PH2 and PH3) in the multicenter study: Institute for Clinical and Experimental Medicine (IKEM) in Prague—1.5 T Siemens Vision (U1), Ninewells Hospital and Medical School in Dundee—1.5 T Siemens Symphony (U2), General Universitario in Madrid—1.5 T Philips Gyroscan (U3), Chambor in Mons—1.0 T Siemens Impact (U4) and 1.5 T Siemens Symphony (U5). A 3D_Flash sequence was chosen as the standard protocol of MR measurement[16] in the head coils with parameters: TR/TE = 20/6 ms, MTX = 128×256, flip angle = 30°, resolution: 2×1 mm per pixel FOV = 256×256×32 mm. There were 32 phase encoding steps along the slice select direction, which represent axial slice thickness 1 mm. We performed two experiments to cover the whole phantom. We used, as a minimum, 55 slices from each imager. Quality control was not done, the adjustment was made automatically. We checked if the automatic procedures could guarantee the consistency of data across platforms. We strictly followed standard clinical procedure. Data were normalized during an evaluation by the MaZda program (see Sec. II C). According to the type of imager, the TE and TR parameters were set up as close as possible to the standard protocol values. T1 and

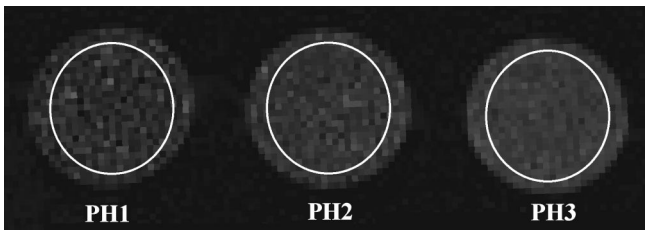

FIG. 2. MR image of PSAG phantoms of different diameters (PH1, PH2, PH3) obtained from imager U5 with marked ROI (circles).

T2 values were obtained from the U1 imager; T2 was measured using a multi-echo sequence with 16 echoes, echo-spacing 22.5 ms, TR=3000 ms, FOV=195×260 mm, MTX=154×256, 3 slices, slice thickness=4 mm, spacing 15 mm. T1 was measured using a set of spin echo sequences, recovery time TR=100, 200, 400, 600, 800, 1000, 1500 ms, echo time TE=22 ms.

The turbo-spin echo sequence was used on the U1 imager to check the influence of resolution to separate different structures (close to the limit of resolution) using TA which we measured with standard sequence parameters (at clinical setup), TR/TE=4700/112 ms, 13 slices, slice thickness =1 mm, spacing 15 mm. We used a resolution of 0.43 ×0.45, 0.77×0.90 and 1.53×1.80 mm per pixel.

C. Data evaluation

Twelve MR slices from each phantom were used for the data evaluation of experiments performed with 4.7 T and 7 T experimental imagers. More than fifty slices from each phantom were used for data evaluation in the multi-center study. The chosen MR images were transferred to a PC and used as input data for the MaZda program for TA.[17] This program allows the computation of a variety of statistical parameters (almost 300) that are derived from a image histogram, absolute gradient, run-length matrix, co-occurrence matrix, autoregressive model and wavelet analysis. We used a circular region of interest (ROI) with a minimum of 300 pixels for a statistical interpretation (Fig. 2).

D. Parameter selection

The texture is better described by several TA parameters than only by one.[3] Therefore we always used sets of ten selected TA parameters—one set was selected manually and two were selected automatically.

The automatic selection of the parameters in the long-term and multi-center studies was performed according to statistical techniques based on the value of the F-coefficient and POE.[4,5,17] Each class represented a different diameter of PSAG phantoms.

The F-coefficient is defined as a ratio of between-class variance to within-class variance (1):

$$F=\frac{D}{V}=\frac{\frac{1}{1-\Sigma_{k=1}^{K}P_k^2}\Sigma_{k=1}^{K}\Sigma_{k=1}^{K}P_kP_j(\mu_k-\mu_j)^2}{\Sigma_{k=1}^{K}P_kV_k}, \qquad (1)$$

where D denotes between-class variance, V—within-class variance, μ_i and V_i—mean and variance of class i, P_i—probability of class i (ratio of the number of data samples from class i to the number of all data samples).

POE is defined as a ratio of the number of wrongly classified data to the number of all analyzed data sets. We considered data as wrongly classified if data were ambiguously assigned to a single class.

The manual selection of TA parameters in the long-term stability testing was done by simple t-statistics for $p<0.01$. We selected invariant parameters characterized by a high level of reproducibility (the range of differences of TA parameter values less than 5%) prior to the study from all calculated TA parameters. To exclude invariant TA parameters which did not carry texture information, we compared TA parameters from MR images of two PSAG phantoms with different sphere diameters (PH1 and PH2). For statistical analysis we chose only invariant parameters that had a difference of mean values of the same parameters greater than 30%.

In the multi-center study the procedure of the manual selection of TA parameters was modified. Selection consists of three steps: (1) we chose the group of invariant TA parameters (characterized by confidence level>97%) obtained from each individual measurement from U1, U2, U3, U4 and U5; (2) then we worked only with data obtained from the measurement at U1 (we chose it as reference), we excluded TA parameters whose mean values were not increasing or decreasing in relation to the diameter of spheres; (3) we chose TA parameters which had the greatest difference of mean values (value $F>20.00$ in analysis of one way variance, critical value is $F=3.03$) among three different PSAG phantoms (PH1, PH2, PH3) measured on U1. In this way we arrived a set of TA parameters that fulfilled these conditions. Ten randomly chosen parameters from that set were used for the comparison in the multi-center tests. The probability level used was $p<0.01$.

For the comparison of data obtained in the two experimental systems, we used the same TA parameters as those that were selected in the long-term stability testing. For an investigation of the possibility of the separation of different structures (PH1, PH2, PH3) using TA which were measured with various resolutions, we used the same procedure of manual selection of TA parameters as in the multi-center study.

E. Statistic evaluation

Statistical evaluation was done according to selected TA parameters by the statistical program B11[17] and by standard statistical methods (t-test, analysis of variance). The B11 software investigates data to assess the ability of the parameters to distinguish various texture categories (in our case various sphere diameters). For classification of the MR images we used Linear Discriminant Analysis (LDA)[18] and Raw Data Analysis (RAW). The success of the classification was categorized by the percentage of correctly assigned subjects. Misclassification means that data were assigned to one

TABLE I. The result of separation of the same and different PSAG phantoms according to chosen TA parameters obtained from MR images measured at experimental Bruker 4.7 T. The classification was performed on standardized data by the RAW and LDA methods. The misclassification (in the table listed as percentage) was calculated for each set of chosen TA parameters (Fisher, POE and Manual).

	The same diameter of the spheres (PH1)		The same diameter of the spheres (PH2)		A different diameter of the spheres (PH1, PH2)	
Set of TA parameters	RAW (%)	LDA (%)	RAW (%)	LDA (%)	RAW (%)	LDA (%)
Fisher	75.8	57.9	38.8	29.3	0	0
POE	68.3	65.7	72.0	31.9	0	0
Manual	89.5	77.4	75.8	39.4	0	0

class incorrectly. All data were standardized, the image intensity range of ± 3 sigma was selected. The normalization ± 3 sigma is equivalent to the range $(\mu-3\sigma, \mu+3\sigma)$ where μ is the image mean and σ denotes its standard deviation.

III. RESULTS

A. PSAG phantoms

We developed new types of phantoms for texture analysis of MR images (Fig. 1). The production of phantoms based on polystyrene spheres and agar is cheap and easy. Changing the diameters of the PS spheres can modify the texture and it is possible to prepare many different structures. T1 and T2 are close to tissue values;[19] the T1 value is 1730 ± 100 ms and the T2 value is 88 ± 3 ms (for phantoms used in the multicenter study). The relaxation times can be easily changed by a different number of spheres in agar (for example we achieved $T2 = 180 \pm 7$ ms for another set of PSAG phantoms used for long-term study). If there is need, it is possible to change the contrast by agar contamination ($DyCl_3$).

B. Selection of the TA parameters

The algorithm for the selection of the TA parameters was tested in three experiments.

(a) In the long-term stability test of PSAG phantoms performed with PH1 and PH2 phantoms over one year, ten TA parameters fulfilled the conditions for the manual selection of the TA parameters. The group of selected parameters contained seven contrast parameters for different distances and angles (TA parameters of the 2nd order histogram) and three wavelet parameters of different scales.[17] Sets including manually selected TA parameters had similar classification results as sets selected automatically in the case of the same or different texture evaluations (Table I).

(b) When the classification of different data (PH1 and PH2) measured at 4.7 T and 7 T was compared, sets with the automatically and manually selected TA parameters gave exactly the same results—0% of misclassification (Table I).

(c) In the case of the multi-center study, 48 invariant TA parameters fulfilled the above mentioned conditions of manual selection. They were TA parameters of the 2nd order histogram, principally based on entropy and sum of square. Ten TA parameters were randomly selected from this group of 48 TA parameters. Sets with automatically and manually selected TA parameters gave similar results (Table II).

C. The long-term stability testing of PSAG phantoms

PSAG PH1 and PH2 phantoms were measured for the period of 12 months. Two tests were performed. In the first, we followed up on changes of selected parameters. A t-test ($p < 0.01$) proved that the parameters were constant during the period of twelve months, e.g., the structure of PSAG phantoms did not change. In addition the separation of PH1 and PH2 phantoms was always successful and the difference of values of TA parameters was higher than 30%. In the second test, we studied the results of phantom classification by using LDA and RAW methods. Table I summarizes the results of the application of these statistical methods. The classification results confirmed the stability and good separability of phantoms.

The separation of PH1 and PH2 phantoms measured at 4.7 T and 7 T was done with 100% success for all the used sets of TA parameters.

TABLE II. The result of the separation of PSAG phantoms of the three different diameters in the multi-center study of whole-body systems. The classification was performed on standardized data by the RAW and LDA methods. The misclassification (in the table listed as percentage) was calculated for each set of chosen TA parameters (Fisher, POE and Manual) and for each imager and for data from all imagers. For one case the classification of data from all imagers by the LDA method according to the POE set, the software procedure failed and we were not able to calculate the misclassification due to the floating point error. Another algorithm (Nonlinear Discriminant analysis) can count misclassification (misclassification = 45.3%).

Set of TA parameters	U1 (%)	U2 (%)	U3 (%)	U4 (%)	U5 (%)	All (%)
Fisher RAW	8.0	0.0	10.7	12.6	19.6	40.5
Fisher LDA	13.3	0.6	13.8	13.2	22.9	51.1
POE RAW	6.8	0.0	46.5	30.2	21.6	45.2
POE LDA	16.1	0.6	28.9	21.1	25.1	-
Manual RAW	41.4	39.9	60.4	54.7	49.3	50.3
Manual LDA	45.8	41.5	50.9	55.0	56.7	56.3

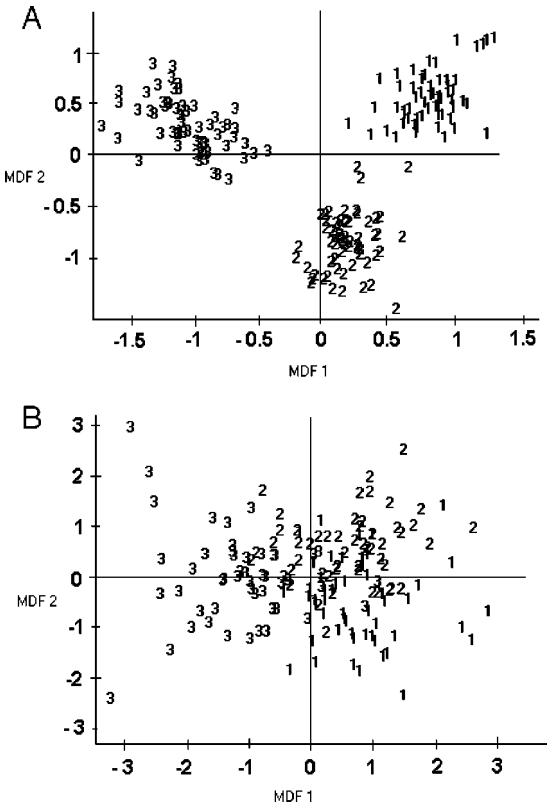

FIG. 3. An example of the separation of PSAG phantoms with different diameters measured with whole-body imagers. Each number represents one slice (number 1: PH3, number 2: PH2, number 3: PH1). The classification was done by LDA methods according to the most discriminating features (MDF1, MDF2) from automatically selected TA parameters [Fisher—(a), POE—(b)]. The excellent separation (0% error) was obtained for data measured in imager U2 (a). The worst result resulted from measurements taken in imager U3, in this case the error was 28.9% (b).

D. The classification of whole body systems

In the multi-center study we compared three PSAG phantoms (PH1, PH2, PH3) in different MR imager systems. We did not obtain the same result of phantoms classification with all MR equipment. The most accurate classification was obtained for data measured with the U2 imager. The worst results of separation of data were obtained from the U3 imager. Figure 3 demonstrates graphical output of good and poor separation of the PSAG phantoms. The separation according to automatically selected TA parameters was better than the manually selected TA parameters. When we added data from all imagers altogether, the average of error increased. Classification by RAW and LDA methods gave similar results (Table II).

E. The influence of resolution on the ability to separate different structures

We tested the influence of resolution on the ability to separate PH1, PH2, and PH3 phantoms according to selected TA parameters by automatic and manual techniques. We achieved no error of separation within resolution of 0.43×0.45 mm per pixel. Within resolution of 0.77×0.90 mm per pixel we achieved only 2.5% of misclassification with the RAW method according to automatically selected TA parameters. The other classification method had no error of separation. The worst results of separation of data were obtained within resolution of 1.53×1.80 mm per pixel. The misclassification with RAW and LDA methods was 2.6–23.1%; with the manual method the misclassification was higher (61.5–64.1%). The results of the influence of resolution to separate different structures (PH1, PH2, PH3) are summarized in Table III.

TABLE III. The result of the influence of resolution to separate different structures (PH1, PH2, PH3) on the U1 imager. The classification was performed on standardized data by the RAW and LDA methods. The misclassification (in the table listed as percentage) was calculated for each set of chosen TA parameters (Fisher, POE and Manual).

Set of TA parameters	Resolution 0.43×0.45 (%)	Resolution 0.77×0.90 (%)	Resolution 1.53×1.80 (%)
Fisher RAW	0.0	2.5	7.7
Fisher LDA	0.0	0.0	2.6
POE RAW	0.0	2.5	23.1
POE LDA	0.0	0.0	7.7
Manual RAW	0.0	0.0	61.5
Manual LDA	0.0	0.0	64.7

IV. DISCUSSION

A. PSAG phantoms

Many different types of phantoms for MR[20] have been developed. Phantoms are necessary to perform standard quality control of imagers and also some are used for tissue structure modeling. Different structures of phantoms should be available with good MR contrast, biochemical and mechanical stability and low cost. To the best of our knowledge only phantoms based on foam were described in the literature[14] for TA. The phantoms with nodular patterns developed in the study are designed for the quality control of TA. They can be also used for easy texture modeling. Phantoms with a stable texture give us an opportunity to optimize the algorithm for the selection of parameters used as input data in statistical classifiers. This even allows a comparison of MR imagers using TA.

B. Selection of the TA parameters

The selection of TA parameters is the most important step of the automatic evaluation of MR images because not all TA parameters carry important image information. It is very difficult to predict which TA parameters can be used for texture classification. The reduction of TA parameters is also necessary from the statistical point of view because a larger set of experimental data is needed for a higher number of TA parameters. We found that techniques for the automatic selec-

tion of the parameters (F, POE) were very sensitive to deviations in pixel interrelationship even to separate the same texture (due to the noise, nonhomogeneity etc., always present in MR images). Thus, they were effective only when we separated different textures.

To overcome this problem, we developed a manual algorithm for the selection of invariant TA parameters easily reproducible for different structures. This selection algorithm based on the invariant parameters is also able to characterize the same structure by TA parameters. According to our results, a manual selection algorithm works well if the resolution is adequate to thr examined texture; in the other case this algorithm fails due to the insufficient resolution (Table III). The choice of invariant parameters seems to be more suitable for the classification of the same textures. The algorithm for selection, which searches the set of parameters automatically, was successfully applied in the case where differences in texture were relatively small and when experimental conditions at different imagers were not quite constant. These two algorithms can be used in various studies in evaluating clinical MR images.[3]

C. The long-term stability testing of PSAG phantoms and classification of whole body systems

We developed the manual algorithm for the selection of TA parameters that could be used for the classification both the same and different textures. The interpretation of misclassification is rather complicated and depends on the character of the examined texture (the same or different). The error of the same texture classification should only relate to data in the order that all data were assigned to one class in the optimal case (at least 50% in case of two classes); for the different texture the error should be 0%. The example of manual selection is long-term testing of differences in PH1 (or PH2) phantoms where we did not find any changes of TA parameters. The case of different textures is demonstrated on the separation of PH1 and PH2 phantoms where we achieved perfect separation using automatically or manually selected TA parameters (see Table I).

In contrast, in the multi-center study of whole body systems we got poor separation (except for automatically selected sets of TA parameters on data from the U1 and U2 imagers) of PH1, PH2 and PH3. An explanation can be found in the insufficient resolution of MR images in comparison to the diameter of PSAG spheres. The separation of different structures is always correct if the resolution (in comparison with the smallest details in texture) is sufficient as is evident from results of separation of different data measured at 4.7 T experimental scanner (resolution $=0.137 \times 0.137$ mm) (Table I) or at 1.5 T whole-body scanner. The resolution 1×2 mm in a multi-center study was chosen in accordance with routinely used sequences for clinical examination. We intended to check the possibility of the separation of different structures (at the resolution limits) using TA which are measured by standard sequence with the standard parameters used in a clinical examination. In critical points where resolution is poor and/or weighted by a large amount of noise (not in our study), the classification of images by human observers is extremely difficult. We tested if the use of TA at these limits (poor resolution) is usable and helpful. Results show that if the sequence does not have sufficient resolution (compare: the resolution of the sequence is 1×2 mm in the multi-center study, PH2 and PH3 phantoms have sphere diameters below the resolution of the MR sequence), the classification according to TA parameters is not accurate. This means that the measurement on different imagers with the same resolution (at the limits of structure detail observation) do not yield the same results due to different technical reasons (magnetic field, type of imager, gradient quality, coils, obsoleteness of scanners and the time of phantom measurements, etc). We were not able to change the protocol of the testing when the tests started. When we tested sites personally, we found slight differences in protocols, which were not identified in previous correspondence. Therefore, the multi-center study represents the real situation at different sites and shows when we can expect good results or not.

To get better results with whole-body systems without modifying the routine measurement protocol, there is a need to use phantoms with sphere diameters larger than the resolution of MR sequence. As is shown in Table III, the ability to separate different structures using TA rapidly increases when we increase the resolution.

The results obtained from the experimental systems confirmed the excellent separation (0% error) of phantoms with sphere diameters much larger than the resolution. In addition to poor resolution of the applied sequence (described above), the manual selection of the reference data could also influence the results because we counted manually selected reference TA parameters only from one imager (U1).

When we compared automatic and manual selection of TA parameters in the multi-center study, we found that automatically selected TA parameters better classify PH1, PH2, and PH3 phantoms. The reason for this is that the automatic selection procedures look for the parameters with the biggest difference from all calculated parameters. Contrary to these automatic procedures, the manual selection of TA parameters is based on prior information obtained from one reference imager (U1).

V. CONCLUSION

The results obtained from the long-term study (measured at 4.7 T) confirm that PSAG phantoms are stable, to at least 12 months, have good separability and reproducibility (verified at 7T) and therefore are suitable for MR imaging. In the multi-center study of whole body systems, the classification of a different texture was influenced by poor resolution which caused large errors. If the resolution is poor, it is of key importance to properly set up the measurement protocol, software and hardware. The resolution is a principal condition to achieve good separation of structure by Texture Analysis procedures, data measured by standard sequence with the standard parameters used in clinical examination is

not able to distinguish different structures (at the resolution limits) by TA evaluation. In this case even powerful statistical methods failed.

ACKNOWLEDGMENTS

We are grateful to Dr. R. A. Lerski, Ninewells Hospital and Medical School in Dundee; Dr. A. Santos and Dr. M. Desco, General Universitario in Madrid; Professor R. N. Muller, Chambor in Mons; and Professor R. Dommisse and Dr. J. V. Audekerke, University of Antwerp, for their assistance with this project. This study was supported by European Commission COST B11; Grant Agency of Czech Republic; Grant No.: 304/03/1189; Ministry of Education, Sports and Youth of the Czech Republic; Grant No.: LN00A065; and Ministry of Health of the Czech Republic; Grant No.: L17/98/23001.

[a)]Electronic mail: daji@medicon.cz

[1] O. Yu, Y. Mauss, I. J. Namer, and J. Chambron, ''Existence of contralateral abnormalities revealed by texture analysis in unilateral intractable hippocampal epilepsy,'' Magn. Reson. Imaging **19**, 1305–1310 (2001).
[2] J. S. DaPonte and P. Sherman, ''Classification of ultrasonic image texture by statistical discriminant analysis of neutral networks,'' Comput. Med. Imaging. Graph **15**, 3–9 (1991).
[3] D. Jirák, M. Dezortová, P. Taimr, and M. Hájek, ''Texture analysis of human liver,'' J. Magn. Reson Imaging **15**, 68–74 (2002).
[4] J. Shurmann, *Pattern Classification* (Wiley, New York, 1996), pp. 264–269.
[5] A. Mucciardi and E. Gose, ''A comparison of seven techniques for choosing subsets of pattern recognition properties,'' IEEE Trans. Comput. **20**, 1023–1031 (1971).
[6] R. Lerski, K. Straughan, L. Schad, D. Boyce, S. Bluml, and I. Zuna, ''MR image texture analysis—An approach to tissue characterisation,'' Magn. Reson. Imaging **11**, 873–887 (1993).
[7] H. Zhu, B. G. Goodyear, M. L. Lauzon, R. A. Brown, G. S. Mayer, A. G. Law, L. Mansinha, and J. R. Mitchell, ''A new local multi-scale Fourier analysis for medical imaging,'' Med. Phys. **30**, 1134–1141 (2003).
[8] S. Murata, K. Mochizuki, T. Nakazawa, T. Kondo, N. Nakamura, H. Yamashita, Y. Urata, T. Ashihara, and R. Katoh, ''Morphological abstraction of thyroid tumor cell nuclei using morphometry with factor analysis,'' Microsc. Res. Tech. **61**, 457–462 (Aug 2003).
[9] A. J. Maubon, J. M. Ferru, V. Berger, M. C. Soulage, M. DeGraef, P. Aubas, P. Coupeau, E. Dumont, and J. P. Rouanet, ''Effect of field strength on MR images: comparison of the same subject at 0.5, 1.0, and 1.5 T,'' Radiographics **19**, 1057–1067 (1999).
[10] B. K. Rutt and D. H. Lee, ''The impact of field strength on image quality in MRI,'' J. Magn. Reson Imaging **6**, pp. 57–62 (Jan.-Feb. 1996).
[11] J. E. Gray and J. P. Felmlee, ''Section thickness and contiguity phantom for MR imaging,'' Radiology **164**, 193–197 (1987).
[12] U. Fischer, W. Doler, S. Luftner-Nagel, and E. Grabbe, ''A multipurpose phantom for quality assurance of contrast-enhanced MR imaging of the breast,'' Eur. J. Radiol. **10**, 162–166 (2000).
[13] F. Podo, ''Quality control in magnetic resonance for clinical use,'' *Ann Ist Super Sanita*, 1994, Vol. 30, pp. 123–137.
[14] R. A. Lerski, L. R. Schad, ''The use of reticulated foam in texture test objects for magnetic resonance imaging,'' Magn. Reson. Imaging **16**, 1139–1144 (1998).
[15] D. Jirák, M. Hájek, and V. Herynek, ''New type of phantom for texture analysis based on polystyrene and agar gel,'' Magma **11**, S343 (2000).
[16] A. Lundervold, COST B11 reports, November, 2000, available: http://www.uib.no/costb11
[17] A. Materka, M. Strzelecky, and P. Szczypinski, Texture Analysis Software MaZda, December, 2000, available: http://www.eletel.p.lodz.pl/cost/progr_mazda.html-MaZda user's manual.
[18] D. L. Swets and J. Weng, ''Using discriminant eigenfeatures for image retrieval,'' IEEE Trans. Pattern Anal. Mach. Intell. **18**, 831–836 (1996).
[19] R. Mathur-De Vre, R. Grimee, F. Parmentier, and J. Binet, ''The use of agar gel as a basic reference material for calibrating relaxation times and imaging parameters,'' Magn. Reson. Med. **2**, 176–179 (1985).
[20] W. M. M. J. Bovée, S. F. Keevil, M. O. Leach, and F. Podo, ''EEC concerted research project: Quality assessment in in vivo NMR spectroscopy. II. A protocol for quality assessment,'' Magn. Reson. Imaging **13**, 123–129 (1995).