The influence of field strength and different clinical breast MRI protocols on the outcome of texture analysis using foam phantoms

S. Waugh[1,2], R. Lerski[1,2], L. Bidaut[2], and A. Thompson[2,3]

[1]Medical Physics, Ninewells Hospital, Dundee, Angus, United Kingdom, [2]University of Dundee, Dundee, Angus, United Kingdom, [3]Department of Surgery, Ninewells Hospital, United Kingdom

Introduction and Aims: Texture analysis in magnetic resonance imaging (MRI), which statistically evaluates grey-level pixel distribution in images, has been used successfully in medical diagnostics [1]. In particular, it has proved useful in Breast MRI for differentiating between benign and malignant lesions and between different types of malignancy [2, 3], and thus has been able to improve the specificity of the examination. It has been reported that texture analysis results are not consistent across clinical scanners, most likely due to differences in acquisition techniques, processing steps and hardware architecture [4]. It has also been suggested that the signal to noise ratio (SNR) is likely to affect the outcome of texture analysis [1].

The aim of this study was therefore to assess - ahead of a new patient study - the effect of two different field strengths and different imaging protocols on the robustness and outcome of texture analysis. Two parameter-matched protocols were run on scanners with different field strengths, and a third protocol was run with a high temporal resolution protocol on the higher field scanner. For each protocol, sequence parameters were altered to assess the impact of changing these on the outcome of texture analysis.

Methods: A breast-mimicking phantom was constructed using lard as a fat substitute and gadolinium-doped agarose gel to represent fibroglandular tissue with T_1 and T_2 parameters that were matched to those obtained clinically. Four different grades of reticulated foam were used as texture phantoms- 90, 75, 45 and 30 pores per inch (ppi) (Foam Engineers Ltd.; Buckinghamshire, UK). These were embedded in the agarose and repeated compression was used to eliminate air bubbles trapped in the foam. The phantom was imaged three times: once (P1) on a 1.5 Tesla (T) MRI scanner (Avanto; Siemens, Erlangen) using a 4-channel breast matrix coil, and twice on a 3.0 T MRI Scanner (Trio; Siemens, Erlangen) using a 7-channel open breast biopsy coil. On the 3.0T scanner two protocols were used: one (P2) with a high spatial resolution sequence matched to the 1.5T protocol acquisition parameters, and one (P3) with a high temporal resolution and lower spatial resolution used for pharmokinetc modelling. Table 1 lists the corresponding sequence parameters.

	Protocol 1 (P1)	Protocol 2 (P2)	Protocol 3 (P3)
TR/TE (ms)	3.8/1.24	3.8/1.28	3.5/1.19
ST (mm)	0.83	0.83	2.0
α (º)	6	6	10
FoV (mm)	320×320	320×320	320×320
Matrix	384×384	384×384	256×256
BW (Hz/px)	650	650	560
iPAT	×2	×2	×2
TA (s)	60	61	24

Table 1- *Sequence parameters for each protocol considered*

The phantom was imaged using the standard sequence (P1, P2 or P3) and then sequence parameters were changed in turn to identify their impact on the outcome of texture analysis. Three different parameters were considered: repetition time (TR), bandwidth/echo time (BW/TE) and flip angle (α). The bandwidth and echo time were linked together by minimising the TE for a given bandwidth, as per the manufacturer's recommended protocol. Five different values were chosen for all the sequence parameters of each protocol, as shown in Table 2. In each case only one parameter was changed at any given time, and TR1, BW1 and α1 represent the baseline values.

TR (ms)				BW/TE (Hz/px /ms)				α (º)			
	P1	P2	P3		P1	P2	P3		P1	P2	P3
TR1	3.8	3.8	3.25	BW1	650/1.25	650/1.28	650/1.11	α1	6	6	8
TR2	4.0	4.0	3.5	BW2	590/1.26	590/1.33	560/1.19	α2	5	5	10
TR3	4.25	4.25	3.75	BW3	540/1.29	540/1.55	530/1.22	α3	4	4	12
TR4	4.5	4.5	4.0	BW4	500/1.31	500/1.57	510/1.23	α4	3	3	14
TR5	4.75	4.75	4.25	BW5	470/1.34	470/1.60	470/1.29	α5	2	2	16

Table 2- *Sequence parameter changes for each protocol*

Texture analysis was carried out using MaZda version 4.7 [5]. Circular regions of interest of no less than 300 pixels were placed in each of the four foam phantoms across the ten central slices. Grey level normalisation was carried out using $\mu \pm 3\sigma$ (μ-grey level mean, σ- grey level standard deviation) to minimise the effect of image brightness and contrast on texture analysis outcome.

Texture features were calculated as derived from the auto-regressive model (ARM), co-occurrence matrix (COM), absolute gradient (GRA), run-length matrix (RLM) and wavelet transform (WAV). As the B11 classification package [5] accepts a maximum of 30 input texture features, the COM features were restricted by choosing an inter-pixel distance of 1 (felt to best reflect fine features) and in only two arbitrarily chosen directions (0º and 45º).

Linear discriminant analysis (LDA) was used to calculate most discriminating features, and classification was then performed using the k-nearest neighbour (k-NN) method, with k=1 and the number of incorrectly identified vectors represented by the percentage of misclassified vectors.

Texture analysis was used to distinguish between the four types of foam for each sequence parameter change at each protocol.

Results and Discussion: Each dataset consisted of 40 data points (4 foam porosities, 10 imaging slices). A total of 13 datasets (baseline plus 4 parameter changes) were analysed five times (for each model) at each protocol.

Both the COM and WAV feature models performed consistently over all protocols at each field strength. The average percentage data misclassifications for these two models over all values of a given sequence parameter are presented in Table 3 for each protocol. Results for the COM features are shown graphically in Figure 1 for every individual sequence parameter change in the different protocols.

	P1		P2		P3	
	COM	WAV	COM	WAV	COM	WAV
TR	12.5	0.0	0.0	0.0	1.0	0.0
BW	9.5	0.0	0.0	0.0	10.5	0.0
α	9.5	0.0	0.5	0.0	2.0	0.0

Table 3- *Average percentage of misclassified vectors for each model at each sequence parameter*

The WAV model provided perfect classification between the four foam phantoms for all sequence parameters across all protocols. The COM model resulted in the highest misclassification at 1.5T (P1), but performed better at 3.0T (P2, P3). Between the two 3.0T protocols, the higher spatial resolution sequence appeared to result in a better texture classification when compared to the lower spatial resolution sequence. This proves consistent with previous findings using spin echo sequences, suggesting spatial resolution is an important parameter when considering texture analysis [6, 7].

There appeared to be no systematic pattern to the rates of misclassification for COM parameters. Although we would typically expect higher rates of misclassification for lower SNR images, there appeared to be no correlation between the percentage of misclassified vectors and measured SNR (r^2<0.117).

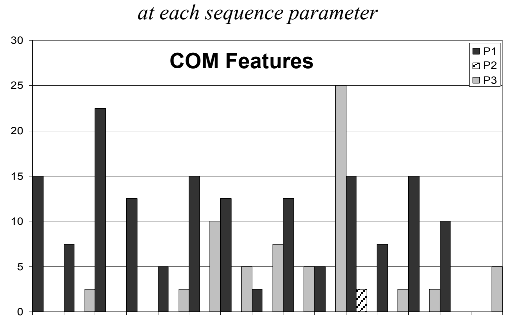

Figure 1- *Percentage of misclassified vectors for each sequence parameter change at three protocols*

Conclusions: Digital texture analysis can reliably distinguish between four grades of foam in a breast phantom despite an identical visual appearance on MR images. The wavelet transform model performed consistently well across all three protocols that were considered for this specific study. The COM features also performed well and even more so at 3.0T and particularly with the high spatial resolution protocol.

References:
[1] Texture Analysis for Magnetic Resonance Imaging, D.M. Hajek M, Materka A, Lerski R. 2006, Med4Publishing: Prague
[2] Gibbs P. Magn Reson Med (2003);50:92-98
[3] Holli K *et al*. Acad Radiol (2010);17:135-141
[4] Collewet G, Mariette F. Magn Reson Imaging (2004);22:81-91
[5] http://www.eletel.p.lodz.pl/mazda
[6] Mayerhoefer *et al*. Med Phys (2009);36:1236-1243
[7] Jirak D, Hajek M. Med Phys (2004);31:616-622