Tina Memo No. 2012-002
Internal

Ice-Water Phantom Localisation for Diffusion Calibration.

Hossein Ragheb, Neil.A.Thacker.

Last updated
20 / 3 / 2012

Imaging Science and Biomedical Engineering Division,
Medical School, University of Manchester,
Stopford Building, Oxford Road,
Manchester, M13 9PT.

Ice-Water Phantom Localisation for Diffusion Calibration.

Abstract

When multiple scanners are used to acquire diffusion MR image data from the same patient (even on the same day), there is no guarantee that identical settings will result in comparable images of a specific organ and region. Despite this, for clinical use, any software developed to compute the apparent diffusion coefficient (ADC) must be expected to give equivalent results. What is needed is a phantom study with appropriate design which supports a calibration, so that appropriate settings for equivalent diffusion measurements can be defined among different scanners. Ideally the calibration process needs to be fully automatic, so that it can be used by non-experts in a clinical setting.

We intend to developed software which automatically locates five cylinders in an ice water phantom (designed for the QuIC-ConCePT project) and then measures the different diffusion values in the cylinders. The location algorithm uses an object recognition process which culminates in robust Likelihood estimation of position and orientation, computed using probabilistic Hough Transforms. This process estimates cylinder locations to an accuracy of a few pixels, even in the presence of field inhomogeneity and significant spatial distortion. Our assessment of performance includes quantification of possible errors due to; data inaccuracy, distortion due to poor shimming, signal to noise, field inhomogeneity and image clutter (i.e. ice). Results indicate that reliable localisation can be obtained using these methods for realistic clinical settings.

Introduction

Diffusion weighted (DW) magnetic resonance (MR) imaging has recently become a focus of attention for facilitating the diagnosis and treatment of tumors [9]. Specifically, to monitor the effect of new drugs administered to treat cancerous organs, DW-MR imaging is an attractive method compared to other alternatives including open surgery. However, there are many difficulties associated with the analysis of DW-MR images [7]. Recomendations have been made for tackling these problems, in order to ensure the validity of any corresponding conclusions. A common quantity that is used when analysing these images is the apparent diffusion coefficient (ADC). For consistency, it is important that any software developed to compute the ADC must give equivalent results on different scanners. One method to achieve this uses a non-biological phantom. For instance, Chenevert et al [2] have recently performed a study using an ice-water calibration phantom and shown that it is possible to get good quantitative agreement between ADC's measured on different scanners. This idea has recently been developed by our collaborators [1] and we have been provided with two diffusion MR data sets. One data set corresponds to the standard shimmed scanner, while the other corresponds to the tune-up shimmed scanner. For diffusion measurement, image data are generally acquired at multiple b values, and exponential fits to this data generate diffusion parameters. In each data set there are five image slices for each of the four b values used, which are 0, 100, 500 and 900 giving a total of 40 images with which to test our software. There are five cylinders of small diameter located inside the main large cylinder which is filled with ice. This maintains the temperature so that the measured diffusion is highly repeatable. One of the five cylinders is placed in the centre while the other four are symmetrically located around the centre (Fig. 1). Each of the five cylinders is filled with a different liquid which has specific diffusion characteristics. For an automated calibration, we must first automatically locate the five cylinders. We intend to use object recognition algorithms to identify and locate the phantom. However, the basic design of the phantom has 4 fold rotational symmetry, and this generates many combinatorial possibilities for location and orientation of the phantom which need to be combined and assessed. Probabilistic Hough transforms are used for this task, implemented as approximations to a likelihood estimation process, in order to deal with the large variations in the constraints on location parameters implied by the measured data.

Work Plan

So far, we have been working with pre-existing software, in order to locate where each cylinder is in the phantom image data. This software requires an edge based shape model, derived from the known physical geometry, so that the circles corresponding to calibration fluids (cylinders) can be located. This model is derived from a simulated image of the phantom using the edge detection and polygonalisation as described below. Specific challenges found

[1] Royal Marsden Hospital

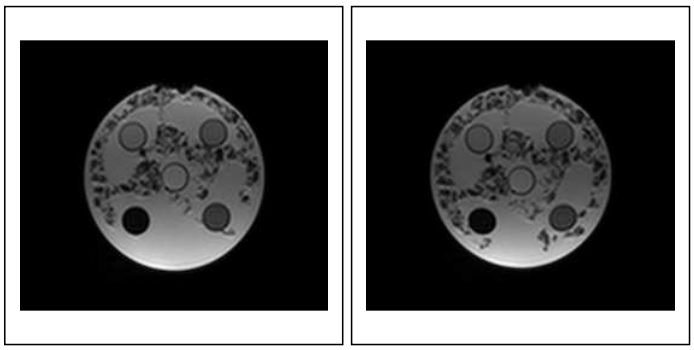

Figure 1: Example original image slices from the standard shimmed (left) and tune-up shimmed (right) phantom data; there are five different liquids with different diffusion characteristics inside the five cylinders; there is ice and water at 0° temperature around the five cylinders where ice appears as black irregularities. The tune-up data show evidence of spatial distortion.

in phantom data have allowed us to identify weaknesses in the original algorithms, which have now been eliminated. We expect to perform more tests and improvements to ensure that results are computed exactly as expected using the underlying statistical assumptions. This document covers the following issues;

- Assessment of our model matching and also the measurements, quantifying how well the model is matched to different image slices. This should identify possible errors due to data inaccuracy including distortion, low signal to noise, field inhomogeneity and image clutter (i.e. ice). We can quantify each of these processes in phantom data and thereby set an allowable range for our software to give reliable results.
- Planned development required for the software to support automated calibration.

Methods and Results

We require a strategy to locate pixels falling inside each of the five cylinders in order to measure the diffusion of the liquid inside each. The larger the number of such pixels, the more accurate the average diffusion measurements, and so the curve fitting. As the measurements from phantom data are pure diffusion measurements, a mono-exponential is expected to be sufficient for curve fitting.

Automatic Location of Cylinders

The existing software operates using a number of separate stages, with the aim of using the observed structure in the phantom image to align to a predefined shape model. Edge data is first extracted from the image and linked into extended curves (Fig. 2.a). These curves are then approximated by polygonal approximations for convenience (Fig. 2.b) [6].

Each line in the polygonal description is then used as the focus for construction of a geometric histogram [5]. Geometric histograms are two dimensional and encode relative orientation and perpendicular distance between lines convolved with a PSF at a scale consistent with these measurements. The histograms can therefore be considered as density estimates of geometric co-occurrence. There are alternative versions of this basic representation which embody varying degrees of invariance. For the current work we used the most discriminative (least invariant) histogram, which considers two possible directions for each scene line separately, in order to compute a full 2π relative orientation. Two histograms are generated for each scene line over a range of possible scales.

The full set of histograms for all lines in the image are matched to a database of similar histograms generated using the model on the basis of a Bhattacharrya overlap [8]. The best match score and scale for each line is interpolated, so that each candidate match provides a constraint on the location, orientation and scale of the phantom in the image. The list of N best candidate matches for each scaled scene line is used as the basis for a series of probabilistic Hough transforms, which compute the robust log likelihood of centre, scale and orientation for the phantom. The 4 fold rotational symmetry of the phantom requires that $N >= 4$, and we have used a value of $N = 8$ in this work. Likelihoods are computed on the basis of a perturbation model for the polygonalisation process, which assumes that errors on lines can be modelled using a constant uniform spherical error distribution

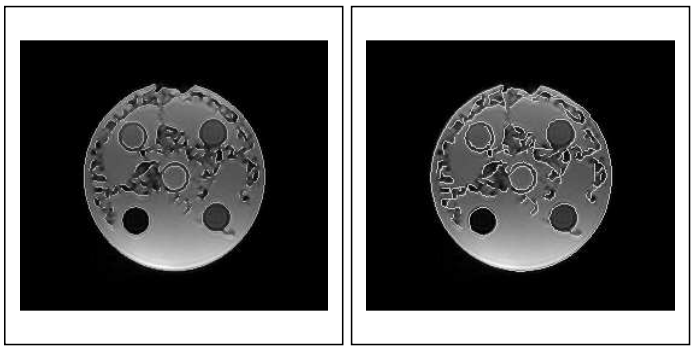

Figure 2: Example image slice is used to show results of canny edge detection (left) and corresponding polygonal approximations (right).

σ_p on the end points [1]. These errors are propagated through to the parameter spaces in order to compute the required Hough entry distributions. The error model assumption is checked and the over-all scale of the spherical error determined using corresponding chi-squared distributions in data. A localisation error model with $\sigma_p = 2$ pixels was found to be approximately the correct value for this data.

Orientation and scale Hough transforms are computed as 1D histograms. The location Hough transform is 2D and entries are made for pairs of matched scene lines (n, m), i.e. a tuple Hough transform. For Gaussian density functions, the equivalent probabilistic form for this Hough transform $H(x, y)$ used to find the centroid of the phantom is given by

$$H(x,y) = \sum_m \sum_n log[p(x,y|n,m)]$$

$$= \sum_m \sum_n log[p(x,y|n)p(x,y|m)] = \sum_m log[p(x,y|m)] \sum_n log[p(x,y|n)]$$

so that the total Hough entry can be considered as the square of the log Likelihood (i.e. standard Hough transform)

$$H(x,y) = L(x,y)^2 = (\sum_m log[p(x,y|m)])^2$$

In practice the 2D Gaussian distribution $p(x, y|n, m)$ is computed directly using error propagation and then truncated at 3 S.D. to form a robust kernel. In comparison to the conventional Hough transform, the tuple based construction helps to remove background from the Hough array (Fig 3.). Cuts are placed on the allowable accuracy of each pairwise constraint, as well as relative orientation in order to improve computational efficiency and further reduce background entries.

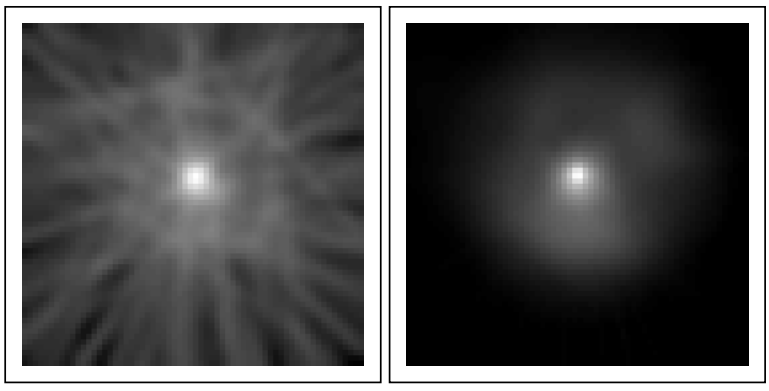

Figure 3: Hough transform data for standard (left) and tuple Hough transfroms (right).

Estimated location, orientation and scale are the result of a global search of the parameter space (followed by a quadratic peak fit), thereby avoiding issues associated with local minima. These parameters are then available to project the model into the image in order to predict the location of each cylinder.

The combined process is expected to generate accurate estimates of phantom position even in the presence of image features such as inhomogeneity and ice, as well as being robust to significant spatial distortion and poor

SNR. These issues would be expected to compromise approaches based upon direct grey-level image matching and iterative optimisation.

Evaluation of Cylinder Localisation

As a simple approach, we assess how well our algorithms can locate the cylinders for all image slices provided across all b values (40 images in total). This can be done by measuring the the largest pixel shift between predicted and observed cylinder edges. Once this value is known it is possible to predict the quantity of usable pixels which might be reliably found in a circle around the predicted cylinder centre. Pixels falling inside these circles, may then be used to compute an average diffusion measurement. Fig. 4 shows the model overlayed on the image slices of Fig. 1. Here for the standard shimmed phantom a good match is achieved while for the tune-up shimmed there is a slight mis-match. Fig 5 shows a systematic clockwise rotation of the model relative to the data. As the centroid of the phantom is easily estimated, the main localisation error for undistorted image geometry is generally of this nature, allowing us to attribute a sign (clockwise or anti-clockwise) to the residual error.

Tabulated results (Table 1) show that the mean shift of the model is comparable for both data sets, though there is evidence of a systematic rotation for the tune-up shimmed data, due to significant systematic spatial distortion (columns 2-3). In both cases however, the mean maximum observed shift is ≈ 3.5 pixels for an internal cylinder diameter of 18 pixels (i.e. 20%). This error does not seem to come from problems of statistical stability (which is expected to be much better than one pixel), but rather spatial distortion of the original image.

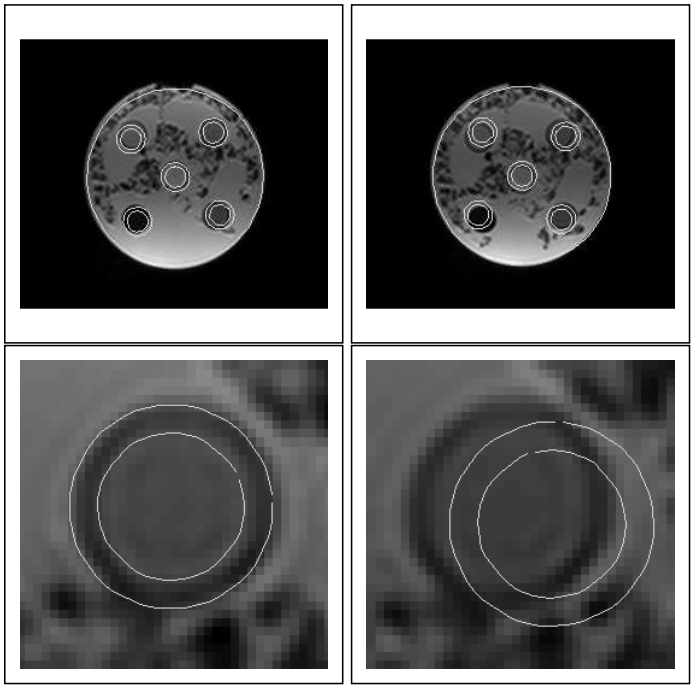

Figure 4: Locating all cylinders on image slices from standard shimmed (left) and tune-up shimmed (right) data (Fig. 1) using our existing software; the display tv before zooming (top) and after zooming in to count the maximum out-of-fit pixels (bottom).

In Table 1 (columns 4-5), we also list the percentage of field inhomogeneity based on the maximum coil correction (see Fig. 6) together with the mean (μ) and the standard deviation (σ) for same image slices of different b values. The degree of this inhomogeneity is substantial (a factor of two across the phantom) but does not appear to produce significant localisation errors. This is expected as the localisation is based upon edges not absolute grey-level values.

In Table 1, we also list the percentage of ice in each image slices. These values are computed from the $b = 900$ (columns 6-7), via a manually controlled thresholding process following inhomogeneity correction, to an accuracy of around 5%. These typical percentages show that there is no observable correlation between the amount of ice and the inaccuracies seen in model matching (columns 2-3). We notice that these percentages were slightly larger for the case of tune-up shimmed images.

Finally, in Table 1 (columns 8-9), we list the percentage of signal to noise ratios (SNR). Each ratio is computed from the distribution of second derivatives (for x and y) around zero, in a central rectangular region including all

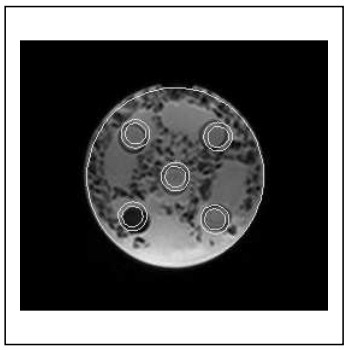

Figure 5: The main error in localisation is generally in the form of a rotation. The maximum observable rotation error can be given a sign (positive if clockwise) in order to help identify systematic rotations across multiple images.

five cylinders, following coil correction. It is clear from the table that the SNR percentages are larger for the case of tune-up shimmed data. However, it can be seen that for different b values the SNR changes by amounts which are small compared to measurement accuracy. Consequently, this level of noise is not large enough to test the robustness of the software to poor SNR. Hence, we selected an image with 50% ice and low geometrical distortion and added extra random noise until the software failed to locate the cylinders (the midpoint of the predicted circles lies outside the required liquid). It was found that for a factor of 2.5× more noise there is no observable degradation in performance, while the localisation algorithm fails 50% of the time at 3.4×, at which point the edge detector begins to fail to extract phantom structure. These noise levels are well beyond the expected levels for phantom data, consequently poor SNR is not expected to play a large part in the failure of localisation.

Given the high degree of inhomogeneity(due to both receiver sensitivity and possibly also B field inhomogeneity) in these images, it is worth considering the possibility that the observed signal does not scale as a function of b in the expected manner. A true diffusion process should scale the observed image data by a constant value (dependent upon the diffusion parameter). In Fig. 6, we investigate whether diffusion is consistent in the first image slices of the shimmed image corresponding to $b = 100$ and $b = 900$. We scale the diffusion values so that the intensity of majority of pixels in the two images agree. Diffusion data are then plotted across the phantom in an identical vertical direction in the middle of each image slice. It is clear from the plot that there are slight differences ($\approx 10\%$) in the diffusion images of water at the top and bottom of the phantom. These differences are consistent across all slices of both datasets and in each case it is the b 900 image which has less signal than expected at the bottom of the phantom.

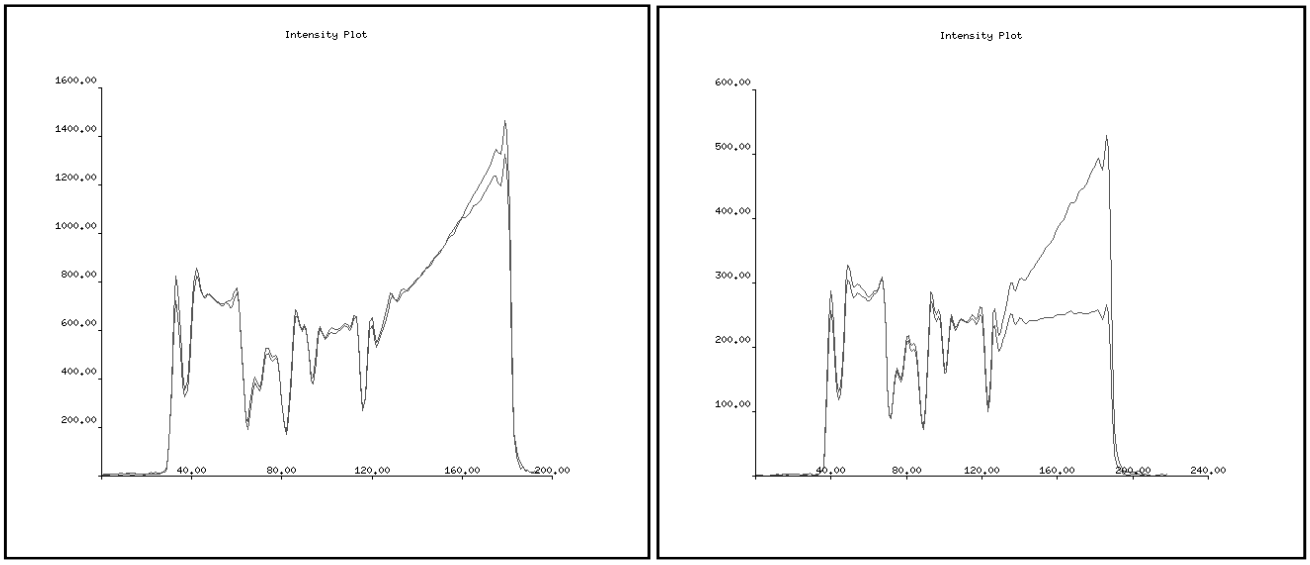

Figure 6: Diffusion data in the phantom across a vertical direction in the middle of (a) two corresponding image slices (first slices in two volumes) of different b values (100 and 900); (b) first image slice corresponding to $b = 900$ for the original image (peaked distribution) and after applying coil correction (flat distribution). Maximum degree of percentage correction (as quoted in Table 1) is taken as the maximum ratio between the coil-corrected and uncorrected values.

image slices(b)	shift (std)	shift (tnp)	inhomogeneity (std) %	inhomogeneity (tnp) %	ice (std) %	ice (tnp) %	SNR (std) %	SNR (tnp) %
01(0)	+2	+3	100	140			1.45	1.76
06(100)	+2	+2	142	129			1.69	1.75
11(500)	+3	+4	125	137			1.41	1.66
16(900)	+2	+2	100	120	51.8	60.1	1.58	1.96
			μ =116.7, σ =20.54	μ =131.5, σ =8.96				
02(0)	-1	+2	100	143			1.66	1.93
07(100)	+4	+3	108	121			1.47	1.79
12(500)	+2	+2	100	100			1.59	1.75
17(900)	+3	+2	100	120	54.6	61.9	1.69	2.01
			μ =102, σ =4.0	μ =121, σ =17.57				
03(0)	+5	+2	76	106			1.70	2.20
08(100)	+2	+4	125	114			1.74	2.39
13(500)	-3	+4	100	100			1.67	2.00
18(900)	+1	+5	80	120	52.9	63.0	1.73	2.03
			μ =95.25, σ =22.44	μ =110, σ =8.79				
04(0)	+3	+2	114	120			1.81	1.98
09(100)	+3	+2	93	93			1.59	1.98
14(500)	-3	+4	78	80			1.74	2.27
19(900)	+3	+1	80	83	55.2	65.3	1.87	1.94
			μ =91.25, σ =16.56	μ =94, σ =18.2				
05(0)	-2	+3	100	100			1.81	2.20
10(100)	-2	+2	80	87			1.66	1.80
15(500)	+4	+2	78	80			1.83	1.95
20(900)	-3	+1	80	100	63.6	65.8	1.93	2.07
			μ =84.5, σ =10.38	μ =91.75, σ =9.94				

Table 1: As the largest source of error is due to a rotation, the maximum shift is diagonal (columns 2-3) at the location of the cylinders (multiply by $\sqrt{2}$ for pixels); the positive sign corresponds to the clockwise rotation while the negative sign corresponds to the anti-clockwise rotation; percentage of field inhomogeneity based on the maximum coil correction (columns 4-5); the mean (μ) and the standard deviation (σ =) are computed for same image slices of different b values; percentage of ice in five image slices corresponding to b = 900 (columns 6-7); percentage of signal to noise (columns 8-9); results are listed for both the standard shimmed (std) and the tune-up shimmed (tnp) image data sets.

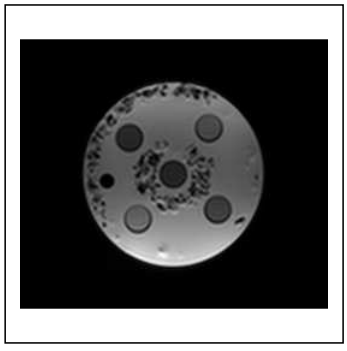

Figure 7: A sample image slice (b = 100) from the new diffusion data where a marker has been added to determine the orientation of the phantom.

Conclusions

Our results suggest that the edge-based localisation method we have developed is robust to the main issues determining phantom data quality, such as field inhomogeneity, ice and SNR. In order to stay within the same range of image quality as seen in our tests we recommend that the quantity of ice is not allowed to rise above 70% and the SNR in the water to be kept above 5%. Inhomogeneity should not be allowed to rise above 150% in order

to remain consistent with this study. Although no particular problem with localisation is expected with increased inhomogeneity, we might start to see compromised fitting of data within the cylinders.

We have seen no evidence that the localisation process will not find the phantom in the 40 images provided with the standard parameters used for testing. Since there are five image slices per b value, it would be possible to check which of the five centres statistically agree (majority vote) in order to remove occasional localisation failures. Even assuming a 1/40 failure rate, the overall success rate would then be expected to exceed 99% for data which conforms to the above requirements.

It should be noted that the fixed value of SNR for varying b values, seen here in water, was unexpected. Closer insepction seemed to indicate a source of structured noise in water due to ringing artifacts [4]. In this case such effects would be less noticable in clinical data.

The signal difference we observe at the bottom of the phantom could be explained by temperature differences, but would require a difference of ≈ 2 degrees centigrade which is considered unlikely (see comments below). It is more likely that this is a change in effective 'b' value, which has been observed to vary spatially and between vendors [3]. The need to calibrate the spatial dependency of an effective 'b' value is perhaps beyond the scope of a clinical setting. So we must either find a way of minimising this effect in data or assume for now that this issue will be addressed by the vendors (see comments below).

Future work now requires us to address two main issues; the main cause of localisation error is the spatial distortion which arises due to poor shimming. This cannot be accommodated by a localisation algorithm which assumes a rigid phantom model and implies that something must be done to accommodate slight shifts during the estimation of diffusion parameters. Also, the current algorithms localise the phantom with a 4-fold rotational ambiguity, due to the symmetry of the phantom. An additional marker has now been added to the phantom which we intend to use to select the correct rotation and as a quality check on predicted location (Fig. 7).

The current evaluation was carried out using available software with variable levels of coding quality, this will need to be modified to conform to GCP and then re-tested. This work will be followed shortly by development of the curve fitting software and the production of a stand alone version of the software which receives image data and delivers measurements. The core of the software will run from a Linux command with specific input and output parameters and is intended to be run as required on the QuIC-ConCePT server.

Comments from D. Collins (*our response*)

I have not found a consistent variation in temperature, certainly not a 'bias field effect'. An ADC map should provide evidence one way or the other. At an NCI meeting about a year ago Tom Chenevert showed that the b-value varies spatially on all MRI scanners and is related directly to the gradient linearity. The b-value variations are greater the further you are away from isocentre and quadratic along z. This almost certainly what you are observing. We still await Tom's paper so the best we can do is report a personal communication from Tom. The vendors agreed that they would try and perform the correction for accurate b-value specification.

We have modified our conlusions to more specifically exclude temperature as an explanation for the drop in signal at b 900 at the bottom of the phantom, we concurr with this explanation and have included the citation.

The SNR scaling is interesting I have yet to give the issue sufficient thought. The noise distributions are not Rician, we have evaluated a number of candidate distributions Gamma, inverse gamma, Chi... depending on the experimental conditions we have observed different distributions. I agree Rician is what we would expect put in practice it is not.

We agree and have have removed all reference to Rician noise.

I would like to see an additional reference to Delakis PMB who developed a sucrose phantom.

This paper also seems to suggest that the best explanation for what we see w.r.t. SNR is structured noise. We have therefore modified our conclusions accordingly.

References

[1] A.P. Ashbrook, N.A. Thacker, P.I. Rockett and C.I. Brown, Robust Recognition of Scaled Shapes using Pairwise Geometric Histograms, *Proc. British Machine Vision Conference*, pp. 503-512, 1995.

[2] T.L. Chenevert, C.J. Galban, M.K. Ivancevic, S.E. Rohrer, F.J. Londy, T.C. Kwee, C.R. Meyer, T.D. Johnson, A. Rehemtulla and B.D. Ross, Diffusion Coefficient Measurement Using a Temperature-Controlled Fluid for Quality Control in Multicenter Studies, *J. Magnetic Resonance Imaging*, 34:983-987, 2011.

[3] T. Chenevert, Private Communication. 2011.

[4] I. Delakis, E.M. Moore, M. O Leach, J.P. De Wilde, Developing a Quality Control Protocol fo Diffusion Imaging on a Clinical MRI System, Phys. Med. Biol. 49, 1409-1422, 2004.

[5] A.C. Evans, N.A. Thacker and J.E.W. Mayhew, The Use of Geometric Histograms for Model-Based Object Recognition, *Proc. British Machine Vision Conference*, pp. 429-438 , 1993.

[6] D.G. Lowe, Three-dimensional Object Recognition from Two-dimensional Images, *Artificial Intelligence*, vol. 31, no. 3, pp. 355-395, 1987.

[7] A.R. Padhani, G. Liu, D. Mu-Koh, T.L. Chenevert, H.C. Thoeny, T. Takahara, A. Dzik-Jurasz, B.D. Ross, M. Van Cauteren, D. Collins, D.A. Hammoud, G.J.S. Rustin, B. Taouli and P.L. Choyke, Diffusion-Weighted Magnetic Resonance Imaging as a Cancer Biomarker: Consensus and Recommendations, *Neoplasia*, 11(2):102-125, 2009.

[8] N.A.Thacker, F.Ahearne and P.I.Rockett, 'The Bhattacharyya Metric as an Absolute Similarity Measure for Frequency Coded Data.' Kybernetika, 34, 4, 363-368, 1997.

[9] B. Turkbey, O. Aras, N. Karabulut, A. T. Turgut, E. Akpinar, S. Alibek, Y. Pang, S.M. Erturk, R.H. El-Khouli, D.A. Bluemke, P.L. Choyke, Diffusion-weighted MRI for Detecting and Monitoring Cancer: A Review of Current Applications in Body Imaging *Diagn Interv Radiol, Turkish Soc Raiol*, 18(1):46-59, 2012.