Phys. Med. Biol. **50** (2005) 1343–1361 doi:10.1088/0031-9155/50/7/001

A complete distortion correction for MR images: I. Gradient warp correction

Simon J Doran[1], **Liz Charles-Edwards**[2], **Stefan A Reinsberg**[2] **and Martin O Leach**[2]

[1] Department of Physics, University of Surrey, Guildford, Surrey, GU2 7XH, UK
[2] Cancer Research UK Clinical Magnetic Resonance Research Group, Institute of Cancer Research, Downs Road, Sutton, SM2 5PT, Surrey, UK

E-mail: S.Doran@surrey.ac.uk

Received 8 October 2004, in final form 5 January 2005
Published 16 March 2005
Online at stacks.iop.org/PMB/50/1343

Abstract

MR images are known to be distorted because of both gradient nonlinearity and imperfections in the B_0 field, the latter caused either by an imperfect shim or sample-induced distortions. This paper describes in detail a method for correcting the gradient warp distortion, based on a direct field mapping using a custom-built phantom with three orthogonal grids of fluid-filled rods. The key advance of the current work over previous contributions is the large volume of the mapping phantom and the large distortions (>25 mm) corrected, making the method suitable for use with large field of view, extra-cranial images. Experimental measurements on the Siemens AS25 gradient set, as installed on a Siemens Vision scanner, are compared with a theoretical description of the gradient set, based on the manufacturer's spherical harmonic coefficients. It was found that over a volume of $320 \times 200 \times 340$ mm^3 distortions can be successfully mapped to within the voxel resolution of the raw imaging data, whilst outside this volume, correction is still good but some systematic errors are present. The phenomenon of through-plane distortion (also known as 'slice warp') is examined in detail, and the perturbation it causes to the measurements is quantified and corrected. At the very edges of the region of support provided by the phantom, through-plane distortion is extreme and only partially corrected by the present method. Solutions to this problem are discussed. Both phantom and patient data demonstrate the efficacy of the gradient warp correction.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Since the early days of magnetic resonance imaging (MRI), it has been recognized (O'Donnell and Edelstein 1985) that MR images are not necessarily geometrically accurate representations of the objects that are being examined. A faithful spatial encoding in MRI depends on the production of a precisely specified linear variation in magnetic field across the sample during the imaging process and, for various reasons, it is not always possible to achieve this in practice. Discrepancies between the magnetic field experienced by the sample and that which is 'expected' by the reconstruction algorithm lead to image distortions. These can have serious consequences in a medical context if the diagnostic images are used as a basis for therapy (e.g., in surgical guidance or radiotherapy treatment planning where high levels of geometric accuracy are required). Typically, neurosurgery applications might require sub-1 mm accuracy, corresponding to the ∼0.5 mm (Schad *et al* 1992) precision with which the head can be localized for surgery within a stereotactic frame. Extra-cranial radiotherapy applications are a little less demanding, since internal organs may move between diagnostic imaging and treatment. Here, tolerances on imaging accuracy may be of the order of several millimetres.

A considerable body of work was performed during the late 1980s and 1990s, in which the theoretical background of the subject was outlined and methods for correcting the distortion were developed. An excellent explanation of the different sources of error has been provided by Sumanaweera *et al* (1993). One may divide the errors into two types: inhomogeneities in the static magnetic field B_0 and errors in the additional field provided by the magnetic field gradient coils.

Chang and Fitzpatrick (1990) demonstrated a method for the correction of B_0 inhomogeneities based on the acquisition of two images of the same slice, using 'forward' and 'reverse' frequency-encoding gradients. Sumanaweera *et al* (1993, 1995) solved the same problem using a different method, based on phase mapping, which allowed the absolute value of B_0 to be determined. Schad *et al* (1992) discussed a method of measuring the gradient-induced distortions, using specially constructed phantoms appropriate to cranial MRI with a limited field of view (FOV). Bakker *et al* (1992), Moerland *et al* (1995) described a correction method taking both effects into consideration.

The primary application of distortion correction has been for stereotactic surgery and radiosurgery on the brain. Jones (1993) has reviewed the requirements on the diagnostic imaging process for accurate localization, and it is clear that stereotactic marker positions on uncorrected MR images can be unreliable. There have been many studies over the years (e.g., Bednarz *et al* 1999, Bourgeois *et al* 1999, Walton *et al* 1997) and their final consensus appeared to be that errors in cranial imaging can be reduced to the order of the pixel size, typically ±1 mm, provided that one applies the correction techniques with care. Thus, by the late 1990s, it seemed that distortion mapping for head imaging was, essentially, a 'solved problem'.

However, in recent years, the demand for ever faster and stronger gradient systems, together with shorter bore magnets, has resulted in considerably compromised gradient linearity. Wang *et al* (2004) have recently re-investigated the area in the light of these developments.

This paper will address a related, but separate, issue, namely the use of distortion correction in extra-cranial imaging with large fields of view. There are clear applications in the areas of interventional imaging and stereotactic surgery and the topic has recently come to the fore with the development of extended field-of-view imaging methods based on a moving patient table (e.g. for MR angiography (Polzin *et al* 2004)). Here, gradient distortions lead to unwanted image blurring. A further motivation for research into distortion correction, relevant at our institution, is the potential of MRI for the planning of radiotherapy treatments. MRI already has a major role in defining lesion boundaries and is often used in association with x-ray CT, leading to the need for image registration. Whilst this is relatively easy to perform in the head, where the skull provides very convenient landmarks, the situation is more difficult extra-cranially. There are promising indications that treatment planning might be possible using MRI alone (Lee *et al* 2003). However, this is rarely the case at present, at least in part because the problem of distortion is still perceived to be unsolved (Prott *et al* 2000). The subject was reviewed by Fransson *et al* (2001), who outlined the principles underlying distortion correction and assessed the current state of the art. Their conclusion was that further studies on the impact of MR distortions were needed.

This paper and the accompanying one (Reinsberg *et al* 2005) aim to provide the required answers. The current work provides a complete description of our method for correcting the gradient-induced distortions. It represents to our knowledge the most rigorous assessment yet provided of the quality of distortion data provided by a 'rod-type' mapping phantom. The second paper discusses the correction of static field inhomogeneities and chemical shift effects.

2. Methodology

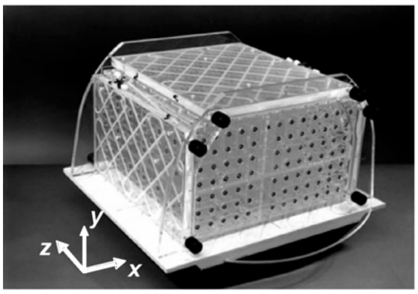
The basic physics and mathematics of the correction procedure are well understood (Chang and Fitzpatrick 1990, O'Donnell and Edelstein 1985). However, a number of issues are either significantly different in the large field-of-view scenario or have not been investigated fully by previous authors. The novel features of this work, previously reported in abstract form in Doran *et al* (2001), are (a) the acquisition of gradient nonlinearity data and its accuracy, with particular reference to the large (>25 mm) corrections required near the edges of the field of view; (b) the creation of 3D distortion maps with an explicit treatment of the through-plane distortion or 'slice-warp' problem; (c) and assessment of the inherent limitations of the 'rod-type' mapping phantom.

A brief overview of the method is as follows: (i) the acquisition of both (distorted) MR and (assumed undistorted) x-ray CT images of a linearity test object; (ii) the identification of corresponding marker positions in the two images; (iii) from a finite number of discrete markers, the production of distortion maps for *x*, *y* and *z* giving distortion values at each voxel; (iv) the 3D correction of patient images using these three maps.

2.1. Rationale and justification

A fact not often made explicit in the gradient-distortion mapping literature is that all the distortion data that one is attempting to measure are, in principle, already known by the scanner manufacturer. The current paths comprising the gradient set are well specified and the relevant electromagnetic software accurate. If a gradient set does not perform according to its specification, it is more likely that this will be because of some gross failure, easily spotted, rather than a subtle shift in the pattern of image distortion.

This begs the question as to the purpose of this and previous attempts to map the distortion field, which are essentially 'reverse-engineering' operations. The scanner manufacturer should be best placed to perform the image correction. The formal justification for this area of research is that (i) a hospital engaged in treatment for which exact geometric position is crucial must have a well-validated quality assurance procedure that goes beyond simply relying on documentation issued by manufacturers and that incorporates empirical verification. (ii) The exact specification and performance of an MRI gradient set are commercially sensitive information and not routinely released to hospitals. (iii) Evidence to date (Wang *et al* 2004) suggests that the correction schemes (predominantly two dimensional) that are applied by the manufacturers are not completely effective in removing the distortion, particularly

(a) 

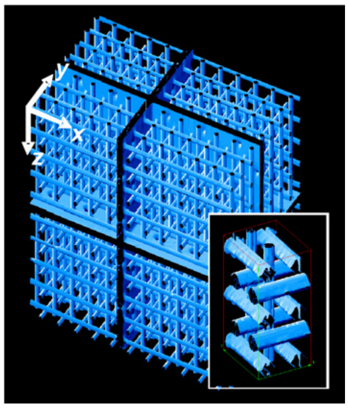
(b) 

Figure 1. (a) Photograph of the linearity test object (LTO) used for generating the distortion mapping data. (b) Rendered view of the 3D x-ray CT dataset used as a reference in the 'spot-matching' process described in the text.

through-plane distortion. (iv) Janke *et al* (2004) state that 'even gradient coils from the same manufacturer will have winding errors and thus variations from the predicted field, such that it would be inaccurate to use a theoretical field expansion.' Part of this work will test the accuracy of the theoretical predictions.

2.2. *Phantom construction*

Details of the linearity test object (LTO) used have previously been published by Tanner *et al* (2000). Its key attributes are (i) a large volume ($440 \times 270 \times 360$ mm^3), with a mapping region of approximately $365 \times 230 \times 340$ mm^3 (height limited by patient couch position); (ii) three sets of orthogonal rods to allow a genuine 3D mapping of the distortion; (iii) full integration into the patient couch in order to obtain a highly reproducible position in the scanner; (iv) a lightweight construction in which a significant fraction of the volume is air; (v) small mapping spots—where extreme distortions occur, a large spot would be highly deformed, making it difficult to obtain an accurate control point. In a refinement of the previously described phantom, extra rows of tubes were added at the top and sides of the phantom to give greater coverage (270 mm and 440 mm respectively) in the anterior–posterior and left-right directions. A small spherical marker was attached at the centre of each face to facilitate registration of CT and MRI data and was used specifically to provide the slice coordinate for matching the central planes in the MRI and CT datasets.

2.3. *X-ray CT scanning*

Prior to MR imaging, the LTO was CT scanned to produce a $256 \times 256 \times 84$ three-dimensional dataset with field of view 500×500 mm^2 in-plane and 5 mm contiguous slices. Particular care was taken in the alignment of the LTO with the physical axes of the CT scanner, and it is estimated that this was achieved within approximately 1 mm. This procedure ensured that the phantom had been constructed to the relevant tolerances, but, more importantly, provided a geometrically accurate spot-matching 'template' against which to compare the MR images—see section 2.5. Figure 1 shows a photograph of the LTO, together with a rendered 3D image based on the CT data.

2.4. MR scanning

All scanning took place on a Siemens Vision 1.5 T scanner and used a 3D gradient echo protocol with the following parameters: TE/TR/α 18.8 ms/5 ms/30°, field of view 480 × 480 mm^2 scanned on a 256 × 256 matrix in-plane, with 84 partitions, each 5 mm, in a 420 mm slab in the out-of-plane dimension. The total image acquisition time was 305 s. This protocol was identical to that used in subsequent patient studies. Since any eddy currents generated are virtually independent of the sample, our assumption was thus that any systematic image distortions produced as a result of eddy currents would be corrected along with those caused by gradient nonlinearities.

Two 3D scans were acquired, identical in all respects except for the sign of the read gradient. The second scan allows for correction of the image distortion in the phantom dataset arising from residual inhomogeneities in the main magnetic field B_0.

2.5. Identification and matching of spots

Distortion data are generated by establishing a one-to-one correspondence between features in the MRI datasets and known physical positions in the phantom obtained from the CT scan. The aim of the first part of the data analysis is thus to identify the distorted positions of the markers as seen on the MR scans. All data analysis was performed on a Sun Sparc Ultra 4, running custom written code in IDL (Research Systems Inc., Boulder, CO) under SunOS 5.8.

In our study, the acquisition used a doubly phase-encoded MR sequence; the data are inherently 3D and can be reformatted for display in any desired orientation. The features to be identified appear as light spots on a 2D grid, where the imaging plane intersects the tubes. The spot positions are measured by finding the centroid of all pixels above a threshold signal in a search vicinity of specified dimension. The problem of locating distortion information in the third, orthogonal dimension will be addressed in section 2.6. The maximum absolute deviation measured was 25.1 mm, whilst using the spherical harmonic expansion of the gradient field—see below—the theoretical value for the absolute distortion for voxels at the very corners of the imaging volume rose to 84.2 mm. There was an upper limit to the distortions that were measurable because some of the spots were shifted completely out of the image. Moreover, for some of the gradient distortions, the spot ‘trajectory’ as one passed down the imaging planes was not monotonic, i.e., going outwards from isocentre, the in-plane position of a distorted spot would move first one way then the other.

These factors meant that establishing the required one-to-one mapping of MRI control points onto genuine physical positions was a non-trivial task. A semi-automated ‘spot-tracking’ algorithm was devised as follows and implemented in IDL.

The inherently 3D CT data were reformatted into, respectively, transverse, coronal and sagittal sets of planes. In each case, it was verified that the pattern of spots remained constant throughout the stack of slices, as would be expected from a set of straight tubes running perpendicular to the imaging plane. From the central plane of the CT scan, a ‘template’ was created manually, defining the location of each spot. Quantitative comparison of this template with the other planes established the geometric fidelity of the LTO, and it will be possible to repeat this at intervals during the lifetime of the LTO to establish that no physical warping has occurred.

Identification of corresponding spots now occurs as follows:

(i) The user is presented with the central planes of the MR and CT datasets side by side and is asked to identify a single pair of corresponding spots as close to isocentre as possible. It is a reasonable assumption that the MR image is negligibly distorted at this ‘tie point’.

(ii) In an automated fashion, the program then looks for a spot in the MR image corresponding to each of the known spots in the CT template. From the positions of the two tie points in the MR and CT images and knowledge of the respective pixel sizes and fields of view, each location in the CT template can be mapped to a corresponding pixel in the MR image. A search is then made in the region surrounding this start point. The search field and thresholds for centroiding were established empirically.

(iii) The program then works outwards from the central slice, matching spots with the CT template for each plane. It is important to note that as spots 'wander', a strategy that simply uses centroiding will often 'converge' on the wrong spot. To avoid this, the start location for the search is based on the distorted positions of the spots identified in the previous plane.

(iv) The operator may correct manually any incorrectly assigned positions, allowing a complete 3D dataset to be processed in all orientations in a few hours. It would be difficult to devise a robust, fully automated solution that could cope well with highly distorted data near the edges of the field of view. Since measuring the gradient distortion is commissioning activity, rather than a frequent QA task, this is a sensible way to proceed.

2.6. Calculation of the distortion maps

At the end of a full spot-matching run, based on two 3D datasets (forward and reverse read gradient), six data files are available, corresponding to a matching operation performed in each of the transverse, coronal and sagittal orientations for both forward and reverse read gradient images. Each file contains the following information: (i) the position of the tie points in CT and MR images; (ii) the coordinates of each spot in the relevant template image; (iii) the coordinates of the spots detected by MR; (iv) a flag saying whether spot matching was successful.

The phantom contains 132 tubes running perpendicular to the coronal imaging planes. The tracking program matched all 132 spots successfully in 126 imaging planes and more than 127 spots in the three other planes in the region of support. The remainder of the planes corresponded either to the walls or exterior of the phantom. The total number of control points in the coronal reformatting of the volume was 16 757. The corresponding figures for the transverse and sagittal planes are 110 spots per plane (7115 control points matched over the whole volume) and 84 spots per plane (16 099 control points matched over the whole volume) respectively. Note that in the transverse reformatting of the original 3D data, planes are separated by 5 mm, corresponding to the original 3D partitions, whereas in the coronal and sagittal reformatting, the plane separation is 1.875 mm, i.e., the pixel size in the originally acquired data. Although the number of points per plane in any given plane is significantly less than that reported by Wang *et al* (2004), the *total* number of control points is significantly greater.

Data from the transverse reformatting of the original 3D image will allow us only to detect distortions in the *xy* plane, whereas a coronal reformatting allows us to detect distortions in the *xz* plane and the sagittal data give *yz* distortions. We thus have the potential for *two* estimates of distortion in each direction, which we label respectively $\Delta\mathtt{x\_tra}$, $\Delta\mathtt{x\_cor}$, $\Delta\mathtt{y\_tra}$, $\Delta\mathtt{y\_sag}$, $\Delta\mathtt{z\_cor}$ and $\Delta\mathtt{z\_sag}$.

The first job is to remove the effect of static field inhomogeneities on the mapping data. This does not require the Chang algorithm (Chang and Fitzpatrick 1990). Since the spots are discrete entities, the 'true' position of the spot is simply the arithmetic average of its position in the forward and reverse gradient images.

Now consider an arbitrary one of the three orientations. The coordinates in which distortion can be measured are labelled x_1 and x_2, whilst the third axis, along which we do not have distortion information, is x_3. The files Δx_tra, etc contain the (x_1, x_2) distortion at each control point, associated with the true (x_1, x_2) position of each point found from CT. The true x_3 coordinate and x_3 distortion are unknown, because of the phenomenon of through-plane distortion. Signal of material from outside the desired plane is distorted into the plane being examined. Equivalently, one can imagine that the image presented is the projection of material from a *curved* or *warped* slice, hence the name 'slice warp' sometimes given to this phenomenon by previous authors. We choose not to use this expression, because it implies that the effect occurs only in multi-slice imaging, whereas, in fact, the phenomenon is also present in true-3D scans where no explicit slice selection occurs in the MR sequence. This phenomenon was illustrated clearly in Wang *et al* (2004).

The user selects the range over which he or she wishes to create a distortion map and the desired grid spacings for the output map in all three dimensions. In addition, the region of support (i.e., the locations where it is believed that reliable distortion information has been obtained) is defined. Under some circumstances this might differ from the desired output region. For example, one might tolerate a small degree of extrapolation outside the measured data points, on the basis that an imperfect correction is better than no correction at all.

The program now loops over all planes in the x_3 dimension for which a sufficient number of control points has been detected. For each plane, a smooth surface is fitted (separately) to the distortion values in the x_1 and x_2 directions, using a minimum-curvature spline (IDL function `min_curve_surf`). This differs from the polynomial model often previously used in distortion analysis (Schad *et al* 1992), exhibiting less extreme divergence outside of the region of support. It is a common interpolation method used for geophysical and remote sensing digital elevation maps. Finally, we note that distortion data might be desired in the out-of-plane direction at different locations than those of the original images and so a cubic spline interpolation is performed in the x_3 dimension.

The problem of through-plane distortion leads to a first-order perturbation in the distortion maps that have been created. Consider our map estimate of the x_1 distortion, $\Delta x_{1\text{est}}(x_1, x_2, x_3)$. In fact, the spot that gave rise to this measurement is actually the cross section through the tube at an incorrect x_3 coordinate, which just *appears* to be in the plane x_3, because of through-plane distortion. This problem arises because the phantom is formed from extended tubes; a phantom consisting of an isolated array of point markers would not be affected in this way. What we need to do is to find the equivalent spot which *would have been* in the image plane had there been no through-plane distortion. This can be found at coordinate $x_3 + \Delta x_3$, where Δx_3 is the x_3 component of distortion. The measured value of Δx_1 should properly be taken as

$$\Delta x_1(x_1, x_2, x_3) = \Delta x_{1\text{est}}(x_1, x_2, x_3 + \Delta x_3) \approx \Delta x_{1\text{est}}(x_1, x_2, x_3) + \Delta x_3 \frac{\partial(\Delta x_{1\text{est}})}{\partial x_3}. \qquad (1)$$

It is clear that this effect is negligible throughout most of the sample volume, since $\Delta x_{1\text{est}}$ changes very slowly with x_3 and also Δx_3 itself is small. However, near the edges of the field-of-view, particularly the corners, this is not so. Our partial solution to this problem is to note that, from a different reformatting of the original data (i.e., one of the other maps that we have created), we have an estimate of Δx_3. Thus, by an interpolation procedure, we are able to calculate the quantity

$$\Delta \tilde{x}_{1\text{est}}(x_1, x_2, x_3) \approx \Delta x_{1\text{est}}(x_1, x_2, x_3 + \Delta x_{3\text{est}}), \qquad (2)$$

which should be a better estimate of the desired distortion. Suppose that we are correcting the map Δx_cor for the effects of through-plane distortion, as in figures 5 and 6. Two estimates $\Delta x_{3\text{est}}$ are available, namely the maps Δ`y_tra` and Δ`y_sag`. Due to the distribution of control points, there are regions where both, one or neither of these maps contains valid data. Correspondingly, the value of $\Delta x_{3\text{est}}$ used in the algorithm is taken as the average of the two values, the value from one of the maps only, or a correction failure is noted.

We may divide the volume for which distortion may be mapped using our phantom into four regions: region I: the gold standard theoretical distortion (see section 2.8) is less than 1 voxel-width and no distortion correction at all is deemed necessary (this applies to 85.84% of voxels in our case); region II: pixels need correction and through-plane distortion does not lead to a significant error in our initial estimate $\Delta x_{1\text{est}}$ (13.57%); region III: through-plane distortion correction is necessary and can be corrected by our algorithm (0.13%); region IV: through-plane distortion cannot be corrected adequately (0.46%).

There are two main reasons for the failure in region IV. Firstly, the effect of through-plane distortion is often to shift spots *outwards* from the isocentre along x_3. In order to measure the true (x_1, x_2) distortion, we would have to measure an x_3 plane further out than is actually available in our original data. Thus, the outer slices in x_3 cannot be corrected using the algorithm. An alternative, not yet implemented, would be to use the measured data to obtain $\partial(\Delta x_{1\text{est}})/\partial x_3$ and use this to *extrapolate* the data. The second reason for failure is that, close to the edge of the field of view, equation (2) does not always provide an improved estimate, because $\Delta x_{3\text{est}}$ may itself be subject to considerable error.

A more robust method, making this procedure iterative and self-consistent is an area of future investigation.

2.7. Correction of the distorted images

Once the measured distortions, $\Delta x(x, y, z)$, $\Delta y(x, y, z)$ and $\Delta z(x, y, z)$, are known, the correction itself is a simple image interpolation procedure. However, when images are distorted, a given volume of material is stretched or squashed into a different volume, thus altering the apparent density and hence brightness in the images. To compensate for this effect, the interpolation must be followed by multiplication by an intensity scale factor, known as the Jacobian.

$$I_{\text{true}}(x, y, z) = J(x, y, z) I_{\text{distorted}}(x + \Delta x(x, y, z), y + \Delta y(x, y, z), z + \Delta y(x, y, z)), \tag{3}$$

where

$$J(x, y, z) = \begin{vmatrix} 1 + \partial[\Delta x]/\partial x & \partial[\Delta x]/\partial y & \partial[\Delta x]/\partial z \\ \partial[\Delta y]/\partial x & 1 + \partial[\Delta y]/\partial y & \partial[\Delta y]/\partial z \\ \partial[\Delta z]/\partial x & \partial[\Delta z]/\partial y & 1 + \partial[\Delta z]/\partial z \end{vmatrix}. \tag{4}$$

In this implementation, partial derivatives of form $\partial[\Delta x]/\partial x$, evaluated at the point (x_i, y_j, z_k), were calculated by finite difference of form $[\Delta x(x_{i+1}, y_j, z_k) - \Delta x(x_{i-1}, y_j, z_k)]/(x_{i+1} - x_{i-1})$.

2.8. Comparison with theory

It is well known that the magnetic field produced by the gradient set of an MRI scanner can be well approximated by a finite set of *spherical harmonics*. The full spherical harmonic expansion of the field is the infinite sum

$$B(r, \theta, \phi) = \sum_{n=0}^{\infty} \sum_{m=0}^{n} \left(\frac{r}{r_0}\right)^n P_{nm}(\cos\theta)[A_{nm}\cos m\phi + B_{nm}\sin m\phi], \tag{5}$$

where P_{nm} is one of the associated Legendre functions (Arfken 1985), r_0 is a scaling constant, A_{nm} and B_{nm} are the spherical harmonic coefficients and (r, θ, ϕ) are the standard spherical polar coordinates. Knowledge of A_{nm} and B_{nm} allows the performance of the gradients to be computed. In principle, an infinite number of such coefficients is required, but, for practical purposes, the behaviour is specified adequately by a relatively small number. A recent publication (Janke *et al* 2004) has used distortion mapping data to generate a set of spherical harmonics as an alternative to mapping the distortion at *all* required points.

The gradient set of the Siemens Vision scanner is the model AS25, which is described by the manufacturers using 29 coefficients for each of the x and y gradients and 7 coefficients for the z-gradient. A calculation using the formula above, with the known spherical harmonic coefficients, was taken to be the 'gold standard' against which to compare all the mapping data generated in this work. Since no information was provided by the manufacturer as to the region of applicability of this theoretical expansion, its validity was tested by the following procedure. (i) The raw image data were distortion corrected using the theoretical spherical harmonic coefficients. (ii) The corrected images were compared with the undistorted CT data via the spot-matching procedure. (iii) Maps of residual distortion were created, with the expectation that, if the gold standard was genuinely correct, these maps should contain only noise.

3. Results

3.1. The measurement process

Figure 2 shows six slices through the unprocessed 3D MRI data from the LTO. For each orientation, the central slice is shown, plus one 'outer' slice at a significant offset from isocentre. The apparent blurring in the vertical direction in figures 2(c)–(f) is due to the low resolution (5 mm) in the original '3D partitions' direction, z. This means that the pixels in images (c)–(f) are non-isotropic. It can be seen that even in the central slices, figures 2(a), (c) and (e), significant distortion is observed at the edges of the field of view. This is particularly noticeable in the coronal image (c). The curved lines at the right- and left-hand sides correspond to two of the straight tubes running in the z-direction. These are visualized as dots in (a) at $x = \pm 220$ mm. Note the aliasing of the tubes on the opposite sides of image (c) above $y = 150$ mm. A particularly striking phenomenon in images (b), (d) and (f) is the appearance of the vertically running tubes in only part of the image. This is caused by through-plane distortion.

Figure 3 is a screen shot from the spot-matching program, stopped at a point where user interaction is required. The program has successfully matched spots for all transverse planes out to $z = -145$ mm. A small number of spots at the edges of the field of view have not been matched correctly. In general, there are three reasons why this might occur:

(i) The spot may be distorted outside the field of view. This imposes a limit on the maximum distortion that can be measured under the current protocol. The easiest solution here would be to enlarge the field of view. However, it was decided not to do this because we wished to use same parameters for the subsequent clinical scans, thereby maintaining the same eddy current conditions.

(ii) The spot moves very close to another feature in the image. This occurs rarely and is easy to correct manually.

(iii) When the spot is highly distorted, its intensity can become very low, making it difficult to detect above the noise threshold. A human operator is sometimes able to detect these low-intensity features where the centroiding software fails.

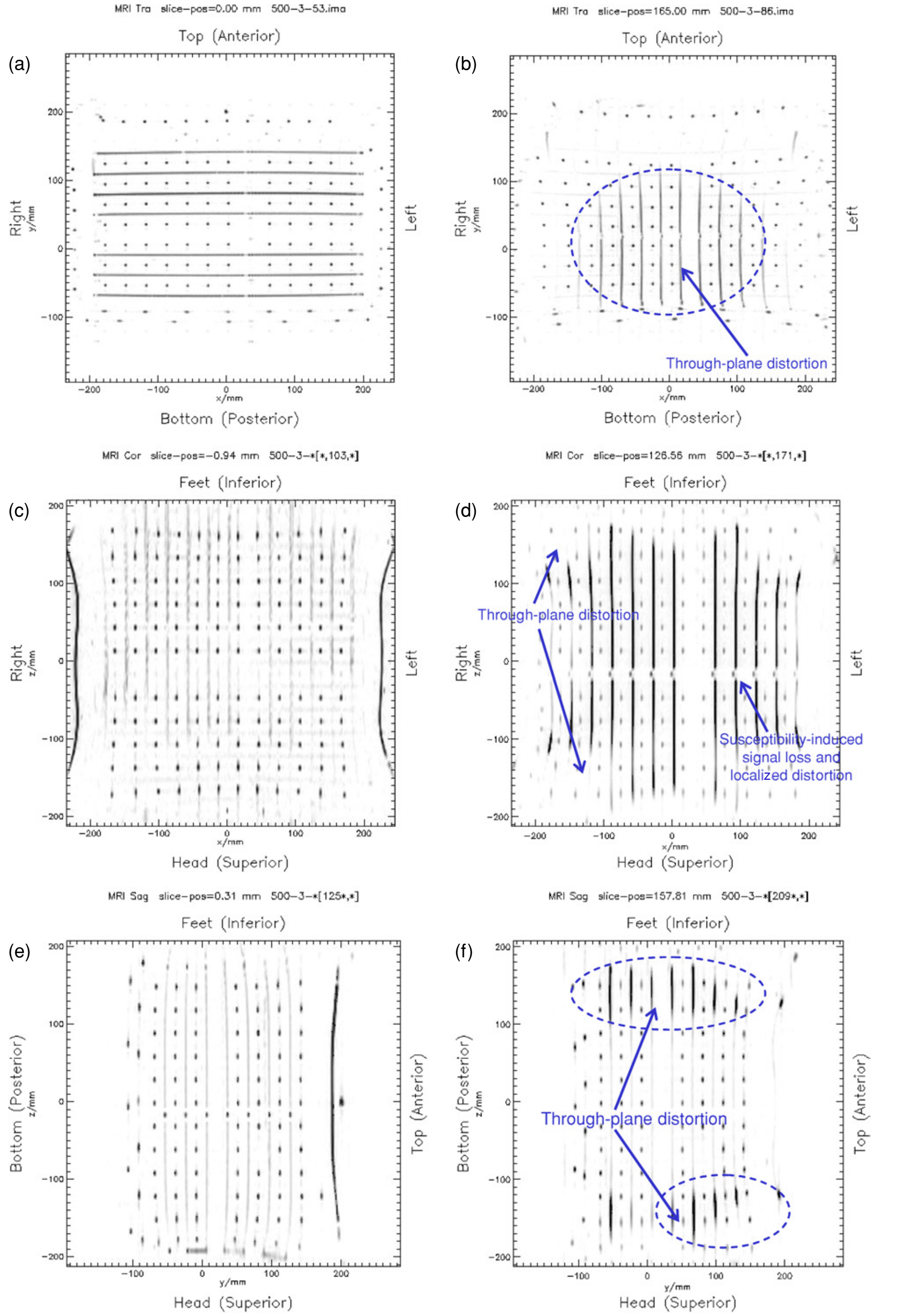


Figure 2. MR images of the distortion mapping phantom, taken from (a, b) transverse, (c, d) coronal and (e, f) sagittal slices through the 3D dataset. Note the severe distortion in the straight tubes running vertically at the edges of the central slices (c, e) and the obvious through-plane distortion in the outer slices (b, d, f). The image colour scales are inverted for clarity.

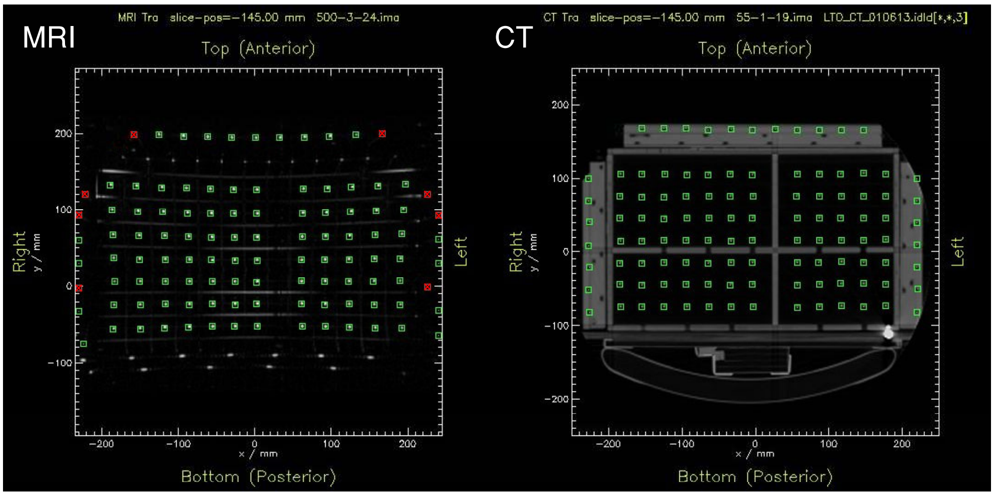


Figure 3. Screen shot of the spot matching process in operation—see main text for further details. Square markers with dots (green) identify spots that have been successfully found and matched, whilst markers with crosses (red) indicate spots that are present in the CT images but which are not successfully matched in the MRI data.

3.2. Region of support

From the design of the LTO, as presented in figures 1–3, it is clear that the region of support (i.e., the volume over which effective mapping data are obtained) does not extend to the edges of the phantom. The extra rows of spots, added to the previous LTO design for this study in the hope of extending the mapping range, are often shifted entirely out of the imaged field of view—see figures 2(a) and 2(b).

Figure 4 gives an overview of this region of support, illustrating the global efficacy of the LTO. For the purposes of this discussion, we will regard the distortion as having been adequately measured by the procedure described in section 2 if the difference between the measured distortion value and the assumed theoretical gold standard is less than one voxel (here 1.875 mm). The 3D field-of-view in each image of figure 4 is $440 \times 270 \times 360$ mm^3, which corresponds to the maximum spatial extent of the mapping tubes. Our distortion maps were generated on an isotropic 2 mm grid, leading to a total of $221 \times 136 \times 181 = 5440\,136$ voxels. In the remainder of the paper, we present here only the results for the x-component of the distortion, but similar figures may easily be generated for the other directions.

Figure 4(a) illustrates the set of points for which a correction is necessary, i.e., for which the gold-standard x-distortion is greater than one pixel. This amounts to 25% of the total. The extent of the problem for large fields of view is revealed by the statistic that 5% of all voxels are distorted by more than 10 mm in the x-direction. Figure 4(b) shows the set of points mapped by our technique as needing correction. It is clear that the form is similar, but that voxels in an outer shell cannot be mapped; the maximum accessible volume is $365 \times 230 \times 340$ mm^3, i.e., about 67% of the total phantom volume. Inside this region, the vast majority of voxels is adequately distortion corrected. Figure 4(c), shows the very small set which have residual distortions of more than one voxel; this constitutes only 1% of the accessible volume. The voxels concerned all lie at the edges of the field of view and in general have large distortions—up to 25 mm. If we relax our agreement criterion to within ± 1.5 voxels, then only 0.06% of voxels remain inadequately corrected.

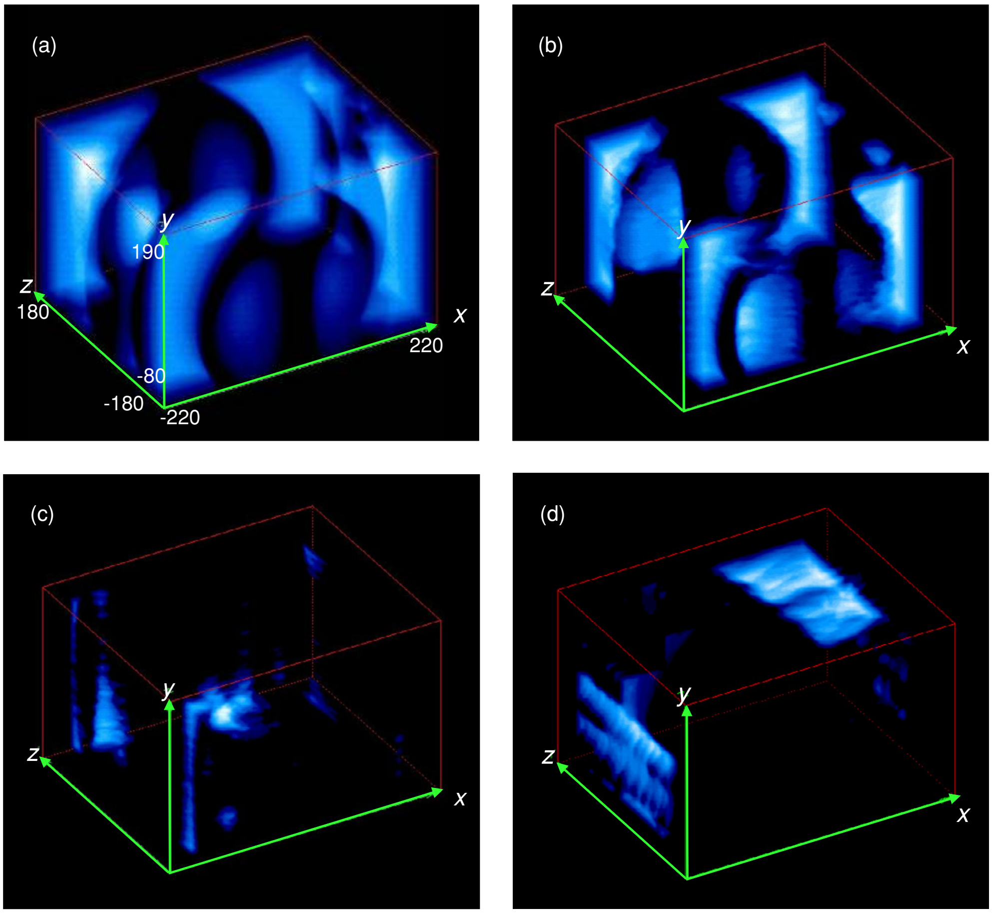


Figure 4. Illustration of the region of support for the LTO mapping phantom: (a) 3D image of the locations of all voxels for which the gold-standard x-distortion is greater than 1 voxel width; (b) all voxels mapped for which the measured x-distortion is greater than 1 voxel-width—note the border round the edge of the phantom that cannot be mapped; (c) all voxels mapped for which the correction is imperfect (i.e., the residual x-distortion after correction is more than 1 voxel); (d) all voxels for which the residual x-distortion after correction with the theoretical gold standard is more than 1 voxel. The images are created by an 'average intensity projection' of a binary dataset and the colour displayed corresponds to the number of voxels along the line of sight.

(Note that these results are slightly worse than expected from looking at region IV in section 2.6. This is because random errors on the spot-matching algorithm occasionally cause very small sub-voxel distortions to be made worse by the correction.)

A very small minority of voxels at extreme distortions may be very poorly corrected. The maximum disagreement between measured and true distortions we observed was 8.7 mm (4.6 voxels). Under these circumstances, it is prudent to define a smaller sub-region in which we can be certain this will not happen. Figures 5 and 6 refer to distortion maps with the following coordinate ranges: $-160\ \text{mm} \leqslant x \leqslant 160\ \text{mm}$; $-70\ \text{mm} \leqslant y \leqslant 130\ \text{mm}$; $-170\ \text{mm} \leqslant z \leqslant 170\ \text{mm}$. Within this volume, 0.6% of voxels were inadequately corrected but there were no voxels at all in the region for which the distortion estimate from our procedure differed from the gold standard by more than 1.5 voxels.

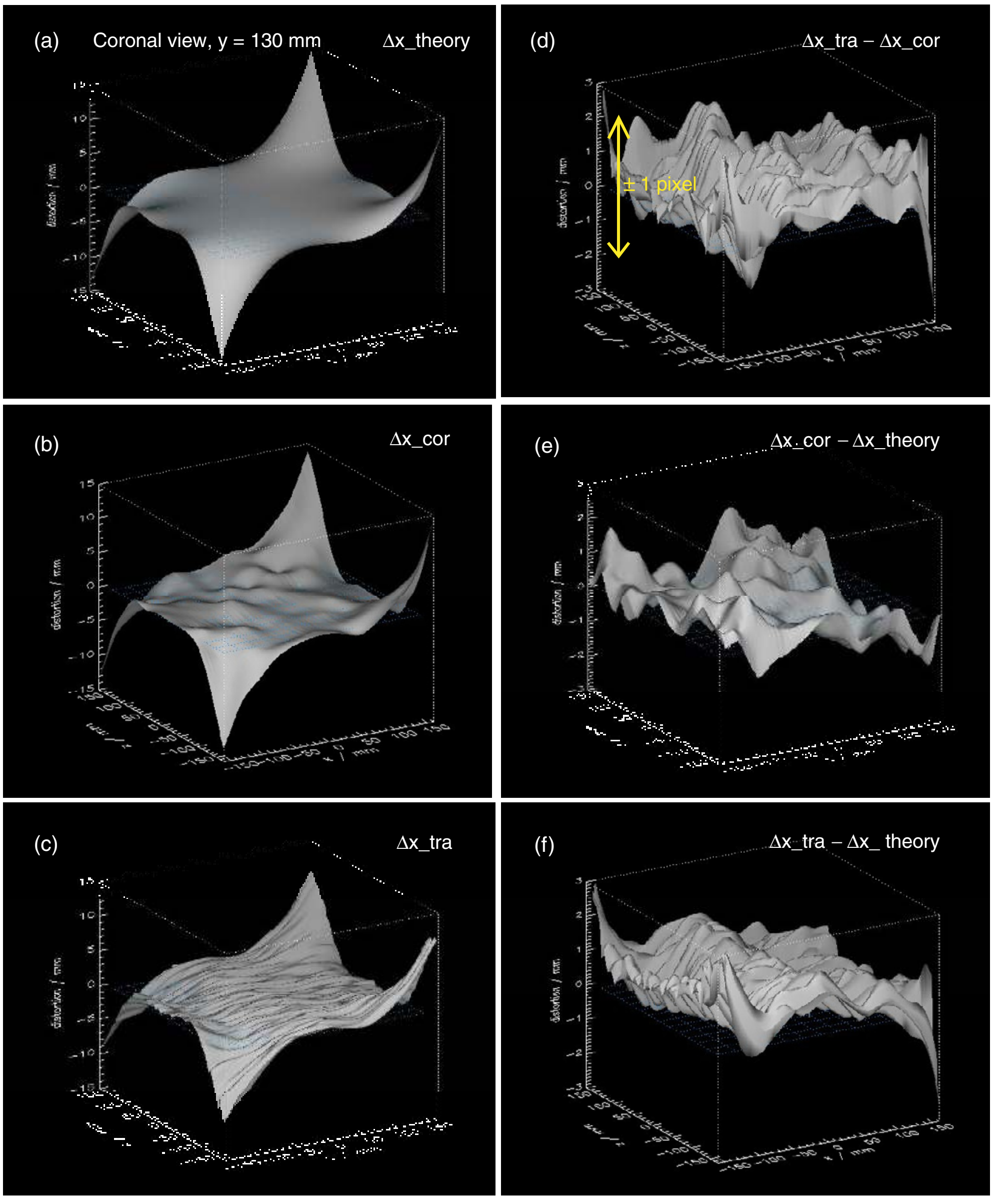


Figure 5. Comparison of experimentally measured and theoretical *x*-distortion data for a typical slice in the coronal direction at an offset of 130 mm from isocentre: (a) theoretical data from spherical harmonic coefficients; (b) map data obtained by running the spot-matching software on the coronal 'reformat' of the 3D image data, i.e., Δ`x_cor`; (c) map data obtained from the raw data viewed in the default transverse format, i.e., Δ`x_tra`; (d) difference between (a) and (b), showing good agreement over the majority of the imaging volume, but significant differences in the corners; (e) difference between (a) and (b), showing some significant differences and a clear trend, but with all deviations between theoretically and experimentally derived measures of distortion less than 1 voxel-width; difference between (a) and (c), showing that the map derived from transverse data is less reliable at the corners.

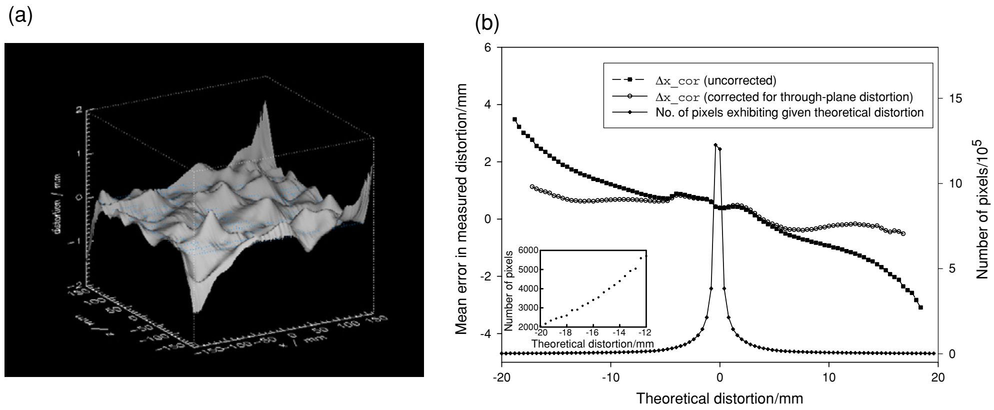


Figure 6. Illustration of the correction of distortion maps for through-plane distortion. (a) Coronal view of the correction to the *x*-distortion map, showing significant values at the corners of the region of support. (b) Left-hand axis: mean error in the measured *x*-distortion for all pixels with a given theoretical *x*-distortion. This plot shows that the errors are not random, but due to the phenomenon of slice warp. The correction applied is successful reducing the mean error well below 1 voxel-width. Right-hand axis: illustration of the significance of the distortion problem in terms of numbers of pixels affected, with (inset) an enlargement for the region of high distortion. The binsize for these histograms is 0.4 mm.

Figure 4(d) gives the results of our attempt to verify the gold standard. All voxels for which the residual *x*-distortion of the corrected images is greater than 1.875 mm are again highlighted, and it will be seen that there are problems at the edges of the volume, although the pattern is different.

3.3. Distortion maps

Figure 5 illustrates the results of the map-generation procedure for a typical plane with moderate distortion (up to 15 mm). As has been described previously, *six* maps are produced: Δx_tra, Δx_cor, Δy_tra, Δy_sag, Δz_cor and Δz_sag. Each of these maps is defined over a full 3D volume. Figure 5(a) presents a 2D coronal slice through the theoretical *x*-distortion map, calculated as described in section 2.8, whilst figures 5(b) and (c) are the corresponding measured results.

Figure 5(d) illustrates the difference between the two independent measurements of the *x*-distortion. It should be noted that the two maps for any given distortion direction (e.g., Δx_tra and Δx_cor) are based on distinct sets of (non-intersecting) tubes, running perpendicularly to each other. The control points are at different places in the images, and it is thus to be expected that slightly different results will be obtained. This is particularly significant at the very corners of the images, where the rate of change of distortion with distance is large.

A particularly noticeable difference between figures 5(b) and 5(c) is the slightly 'corrugated' effect in figure 5(c). Δx_tra is created by fitting a continuous, minimum curvature surface independently to each of a stack of transverse planes. No criterion for continuity of the function Δx *across* planes is defined in the current algorithm. Thus, when we 'slice through' the complete dataset in the perpendicular coronal direction, we see a discontinuous appearance. The magnitude of this effect is approximately 1 mm (half a voxel width). The effect would be seen the opposite way round were we to view *transverse* slices through Δx_tra and Δx_cor.

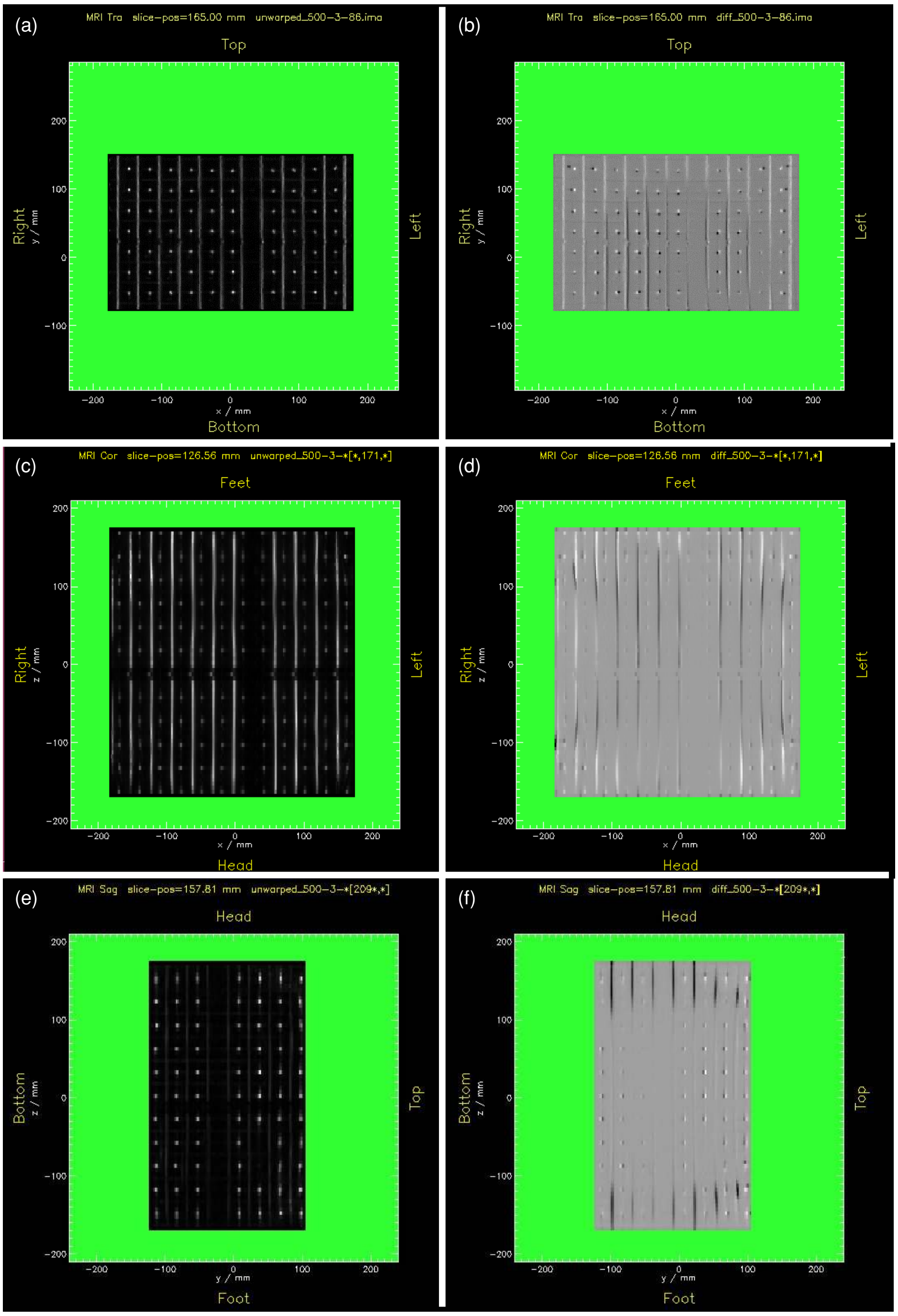


Figure 7. Distortion-corrected versions of figures 2(b), (d) and (f), together with plots of the difference between corrected and original images. Note the excellent rectification of the through-plane distortions and the large shifts of the positions of the spots in the transverse images. The outer, solid border represents regions of the images that are outside the region of support and hence not corrected. Distortion information is available for some of this region, but is not ‘guaranteed’ to be accurate to ± 1 voxel width as it is inside the region.

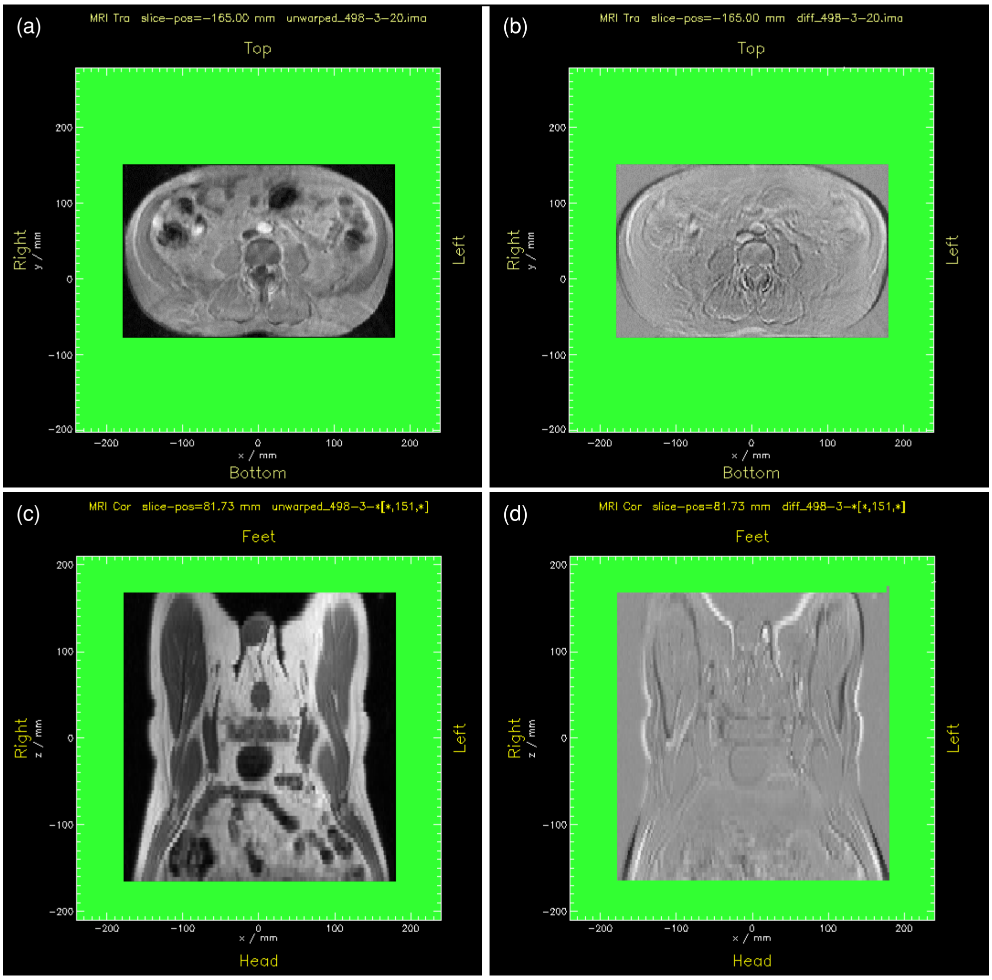


Figure 8. Correction of a typical patient dataset. Note the large distortions at the skin surface. These are potentially important in planning radiotherapy treatments.

Figures 5(e) and (f) show the deviation of the measured distortion from that predicted by theory for maps $\Delta\mathtt{x\_cor}$ and $\Delta\mathtt{x\_tra}$ respectively. It will be seen that over the entire region, $\Delta\mathtt{x\_cor}$ is less than one pixel. At the very corners, though, the error in $\Delta\mathtt{x\_tra}$ rises to approximately 1.5 pixels.

3.4. Through-plane distortion

Figure 6 illustrates the efficacy of the procedure outlined in section 2.6 for dealing with the through-plane distortion problem. Figure 6(a) shows the difference between the initial version of the $\Delta\mathtt{x\_cor}$ map and the map corrected for through-plane distortion effects. For most of the mapping region, the algorithm makes no significant difference to the measured map. However, in the corner regions, where the measured distortion is less than that predicted, the correction serves to 'boost' the value, bringing it much closer to that required. Figure 6(b) displays the data in a different fashion, plotting the mean error in measured x-distortion for all points with a given theoretical x-distortion. There is a clear systematic error, which is significantly reduced when the correction procedure is introduced.

3.5. Distortion-corrected images

Figure 7 shows that the correction algorithm described in section 2.7 works successfully for all planes. The distorted images that have been corrected are figures 2(b), (d) and (f), the outer slices of the LTO, where considerable through-plane distortion is seen. The method does an excellent job of restoring the vertical tubes in figures 7(a) and (c) and removing them from figure 7(e), as well as restoring the spots to a regular grid pattern. The corresponding figures 7(b), (d) and (f) show the difference between the corrected and uncorrected images. The adjacent dark and light spots show clearly where a spot has been moved; this is particularly apparent in the transverse slice figure 7(b). Note how the original data have had a mask applied around the outside of the field of view. This corresponds to the reduced region of support described in section 3.2. Although distortion measurements are available outside this region, inside it, we are able to guarantee the quality of the distortion correction to within one voxel.

Figure 8 shows two views of a 3D clinical dataset, corrected by the method shown here. It is not possible to assess the quality of the correction in this case, since the 'ground truth' is inherently unknown, but merely to note the magnitude of the changes made to the outer portions of the images, as seen in the difference images, figures 8(b) and (d).

4. Discussion and conclusions

Figure 2 demonstrates starkly that for large field-of-view imaging, distortion correction is necessary for applications where accurate geometrical information is required. Our particular interest is in radiotherapy planning, but such corrections have many other uses. The reasons for wanting to measure such distortions empirically rather than relying on data and correction routines from scanner manufacturers have been discussed.

Our method for correcting the images is based on mapping measurements made on a rod-type phantom. The linearity test object differs from previous rod phantoms in that there are three orthogonal inter-penetrating arrays of tubes, which allow a full 3D distortion correction without repositioning the phantom. In order to achieve this, a robust spot identification procedure is required, which has been described in detail.

In the most recent comparable study (Wang *et al* 2004), the MRI features are *dark* crossing points in an xy grid, located in the z-direction by a novel derivative method. The method by which any given MRI feature is related to its genuine position is not described by the authors. However, we note that because of 'serious image artefacts' only a subset of control points defined by the central part of the array was used. This corresponded to a corrected volume of $257 \times 225 \times 257$ mm^3. Within this region, the maximum absolute distortion was found to be of order 9 mm and we must assume that, under such conditions, establishing the necessary one-to-one correspondence between the phantom and the MR image features was straightforward.

To the best of our knowledge, the volume coverage and distortions quantified in this study represent the largest yet discussed in the literature. Mapping data were successfully generated over a region of size is $365 \times 230 \times 340$ mm^3 and, after correction by our method, the x-distortion exceeded one voxel-width for only 1% of voxels and 1.5 voxel widths for only 0.06%. A few voxels at the corners of the volume had residual deviations of up to 8 mm out of a pre-correction distortion of some 25 mm. By restricting the region of support to $320 \times 200 \times 340$ mm^3, x-distortions were successfully mapped to within ± 1 voxel for all but 0.6% and it was possible to guarantee that no voxel had a residual x-distortion of more than 1.5 voxels.

The mean error in measured x-coordinate over the $365 \times 230 \times 340$ mm^3 volume accessible by the phantom is 0.6 mm equal to approximately one third of the voxel width in the original MRI dataset. In our case, the true marker positions were *a priori* unknown and were established using CT images, themselves with a resolution of 1.95 mm in-plane. Although this result appears inferior to that of Wang *et al* (2004) who quote the mean error in the measured coordinates of the control points as being 'of the order of 0.1 mm', we note that our precision is entirely appropriate for the eventual surgical or radiotherapeutic use.

The results above assume that the manufacturer's spherical harmonic description of the gradient set provides an accurate description. An attempt to verify this showed that, at the edges of the field of view, images corrected using these theoretical values still show residual distortion of more than one voxel. However, these data should be treated with some caution. They are based on raw images of limited FOV, and some spots have been distorted completely out of the image. For the same reasons that we are unable to measure experimental distortions over the full extent of the mapping phantom, we are also unable completely to restore the extreme spots for comparison with CT, even if the theoretical values are accurate.

Many of the problems described above could be rectified and results further improved by a modification of the original data acquisition to use a larger field of view. However, this would entail additional studies to determine the impact of using different imaging sequences for distortion mapping and clinical imaging. This work remains to be performed. Further study is also warranted to discover whether reducing the voxel size in the raw data would give a corresponding improvement in the agreement between theoretical and measured distortion values.

An important aim of this project was to investigate the intrinsic limitations of the 'rod-type' phantom. These are evident particularly at the corners of the field of view, where the distortion changes extremely rapidly and we recorded systematic errors of up to 8 mm. As described in section 2.6, it is no surprise that errors in distortion mapping will occur in these regions, because of the phenomenon of through-plane distortion. To avoid confusion, we emphasize the fact that the *image unwarping algorithm* used here, equation (3), *does correct* the direct effects of through-plane distortion, given the correct distortion maps. What remains problematic is to obtain these maps using this type of phantom. An algorithm has been presented that provides a partial solution—see figure 6—but a complete correction is likely to prove impossible. The new design of mapping phantom recently described in Wang *et al* (2004) offers potential advantages, because the dark crossing points occur at well-specified 3D positions and are not extended objects such as the tubes here. Alternatively, a possible extension would be to combine the direct mapping of the distortion fields with the spherical harmonic deconvolution described in Janke *et al* (2004).

However, in all this it is important to bear in mind that the very poorly corrected points are probably merely of academic interest: they lie predominantly outside the anatomy of patients and represent a minuscule fraction of the total volume. In the $365 \times 230 \times 340$ mm^3 volume referred to above, only 0.03% have a residual deviation of more than 2 voxels and 0.003% more than 3 voxels.

Two potential alternatives to the methodology employed in this paper and previous studies might be (a) to use similar methods to those found in cardiac tagging experiments; (b) to perform a non-rigid body registration between the MRI and CT data. However, comparison of the results of these methodologies is beyond the scope of the current paper.

The distortion maps produced have been applied with obvious success to correct the images of the LTO itself and on clinical data. In conclusion, we have developed and verified a reliable method for removing the effects of gradient induced distortion over a large field of view.

Acknowledgments

This work was supported by the UK Department of Health (NEAT grant scheme award B132). SJD is grateful to Siemens Medical Systems for the use of the AS25 gradient spherical harmonic data.

References

Arfken G 1985 *Mathematical Methods for Physicists* 3rd edn (New York: Academic) p 669

Bakker C J G, Moerland M A, Bhagwandien R and Beersma R 1992 Analysis of machine-dependent and object-induced geometric distortion in 2dft MR imaging *Magn. Reson. Imaging* **10** 597–608

Bednarz G, Downes B, Corn B W, Curran W J and Goldman H W 1999 Evaluation of the spatial accuracy of magnetic resonance imaging-based stereotactic target localization for gamma knife radiosurgery of functional disorders *Neurosurgery* **45** 1156–61

Bourgeois G, Magnin M, Morel A, Sartoretti S, Huisman T, Tuncdogan E, Meier D and Jeanmonod D 1999 Accuracy of MRI-guided stereotactic thalamic functional neurosurgery *Neuroradiology* **41** 636–45

Chang H and Fitzpatrick J M 1990 Geometrical image transformation to compensate for MRI distortions *Proc. SPIE: Med. Imaging IV: Image Process.* **1233** 116–27

Doran S J, Moore L and Leach M O 2001 Pelvic MR scans for radiotherapy planning: correction of system and patient-induced distortions *Proc. ISMRM, 9th Annual Meeting (Glasgow)*

Fransson A, Andreo P and Potter R 2001 Aspects of MR image distortions in radiotherapy treatment planning *Strahlenther. Onkol.* **177** 59–73

Janke A, Zhao H, Cowin G J, Galloway G J and Doddrell D M 2004 Use of spherical harmonic deconvolution methods to compensate for nonlinear gradient effects on MRI images *Magn. Reson. Med.* **52** 115–22

Jones A P 1993 Diagnostic imaging as a measuring device for stereotaxic neurosurgery *Physiol. Meas.* **14** 91–112

Lee Y K, Bollet M, Charles-Edwards G, Flower M A, Leach M O, McNair H, Moore E, Rowbottom C and Webb S 2003 Radiotherapy treatment planning of prostate cancer using magnetic resonance imaging alone *Radiother. Oncol.* **66** 203–16

Moerland M A, Beersma R, Bhagwandien R, Wijrdeman H K and Bakker C J G 1995 Analysis and correction of geometric distortions in 1.5-T magnetic-resonance images for use in radiotherapy treatment planning *Phys. Med. Biol.* **40** 1651–64

O'Donnell M and Edelstein W A 1985 NMR imaging in the presence of magnetic-field inhomogeneities and gradient field nonlinearities *Med. Phys.* **12** 20–6

Polzin J A, Kruger D G, Gurr D H, Brittain J H and Riederer S J 2004 Correction for gradient nonlinearity in continuously moving table MR imaging *Magn. Reson. Med.* **52** 181–7

Prott F J, Haverkamp U, Eich H, Resch A, Micke O, Fischedick A R, Willich N and Potter R 2000 Effect of distortions and asymmetry in MR images on radiotherapeutic treatment planning *Int. J. Cancer* **90** 46–50

Reinsberg S, Doran S J, Charles-Edwards L and Leach M O 2005 A complete distortion correction for MR images: II. Mutual information-based rectication of static-field inhomogeneities in MR images *Phys. Med. Biol.* (submitted)

Schad L R, Ehricke H H, Wowra B, Layer G, Engenhart R, Kauczor H U, Zabel H J, Brix G and Lorenz W J 1992 Correction of spatial distortion in magnetic-resonance angiography for radiosurgical treatment planning of cerebral arteriovenous-malformations *Magn. Reson. Imaging* **10** 609–21

Sumanaweera T S, Glover G H, Binford T O and Adler J R 1993 MR susceptibility misregistration correction *IEEE Trans. Med. Imaging* **12** 251–9

Sumanaweera T S, Glover G H, Hemler P F, Vandenelsen P A, Martin D, Adler J R and Napel S 1995 MR geometric distortion correction for improved frame-based stereotaxic target localization accuracy *Magn. Reson. Med.* **34** 106–13

Tanner S F, Finnigan D J, Khoo V S, Mayles P, Dearnaley D P and Leach M O 2000 Radiotherapy planning of the pelvis using distortion corrected MR images: the removal of system distortions *Phys. Med. Biol.* **45** 2117–32

Walton L, Hampshire A, Forster D M C and Kemeny A A 1997 Stereotactic localization with magnetic resonance imaging: A phantom study to compare the accuracy obtained using two-dimensional and three-dimensional data acquisitions *Neurosurgery* **41** 131–7

Wang D, Doddrell D M and Cowin G 2004 A novel phantom and method for comprehensive 3-dimensional measurement and correction of geometric distortion in magnetic resonance imaging *Magn. Reson. Imaging* **22** 529–42