Data Representation: Mapping, Rendering and Visualization

Derek Hill, PhD
Professor of Medical Imaging Sciences
Centre for Medical Image Computing (CMIC)
University College London, London WC1E, 6BT
Derek.Hill@ucl.ac.uk

MRI data provides many challenges for data representation and visualization. The data provided by the scanner has many different formats – 2D, 3D, 4D, scalar valued, complex, vector valued, tensor valued, multi-echo etc. Once the data has been processed, eg: using the sorts of algorithms that have been described earlier in this course, there is an even wider variety of data to represent. This diversity of MRI data has important implications for image analysis and visualization algorithms.

This tutorial begins with some preliminary discussion about the nature of MR data, then describes approaches to visualization. This article does not aim to be comprehensive, but to discuss some of the key issues that are important for effective image analysis and visualization.

Preliminaries

Mapping

A key concept in image analysis is *mapping*. The most important type of mapping in medical image analysis is the geometrical transformation. For image registration and visualization, it is frequently desirable to rotate, translate, zoom, or warp an image. Another important type of geometrical mapping is mapping from the voxel space to an alternative representation of the image. The most common of these is mapping to a plane ("flat map") or a standardized surface shape. These are all mappings defined on a domain, which is likely to be the image field of view, or the intersection of two or more image fields of view. All these types of geometrical mapping are applied to a discrete array of voxel values, not to continuous mathematical functions. Therefore we are mapping intensities within the domain, rather than just positions. For this reason, it is very useful to understand the properties of the intensity information in the images.

Discretisation of data

MRI data are discrete, and any time we apply any geometrical mapping to the images, we need to remember this. An mapping task, which could be a simple rotation, or inflating a brain surface to a sphere, needs to take account of the discrete nature of the images.

The data values exported from the scanner, most often simple scalars, are stored on an array of voxels. The voxel array is a grid that is normally regular and anisotropic. That is, while the voxel spacing is usually constant along a given image axis, the spacing is typically different in x and y (the slice plane) from z (the through-slice direction). This spacing between voxel centres, typically in mm, is often called the voxel size, but it is better to call it voxel spacing for reasons that will become clear below. It is easy to obtain the voxel spacing information from the image headers of just about any medical image format. Using this information, it is possible to view data in the original slice orientation, reformat it in orthogonal, oblique or curved planes, rotate and translate it, and make use of image analysis techniques such as image registration and segmentation. The voxel spacing is, however, only part of the full description of the discrete mature of the images. If we are worried about preservation of quantitative information in the images, we need to take account of other information about the images.

The first useful distinction is between the voxel spacing, and the image resolution. It is very common for the image resolution to be quite different from the voxel spacing, and, whereas voxel spacing is usually the same in both the x and y directions, it is common for resolution to be different between these. Modern MR scanners allow the acquisition matrix and reconstructed matrix to be selected independently. The scanner zero fills prior to re-construction to generate reconstructed voxel spacing smaller than the acquired resolution elements. For a 2D multislice sequence (such as most spin echo or echo planar images), this flexibility is only available in the slice plane. For a 3D volume scan, the reconstructed voxel sizes can also differ in z, but it is common for there to be less flexibility in this direction, with zero filling by a factor of 2 often being the only option available on the scanner console. It is because of this difference between acquired and reconstructed voxel sizes that the term “voxel spacing” is preferable to “voxel size”.

An important issue related to voxel spacing is the point spread function of the image. For an image reconstructed by a Fourier transform, this is a sinc function (sinx/x) which has infinite extent. This means that the value of a particular voxel is not independent of its neighbours. Images with sinc point spread functions are easy to interpolate. The most correct interpolation is sinc interpolation (or equivalent), with linear interpolation (the most widely used sort) producing blurring. However, if the images have already been interpolated onto a larger image matrix by zero filling as part of the reconstruction, then simpler interpolation can be quite sufficient for most purposes. . With a multislice acquisition, the point spread function through the slice-plane is related to the profile of the slice selection pulse. The slice selection pulse is often a truncated sinc pulse, which is designed to give a rectangular slice profile. For efficient multislice imaging, it is desirable for each slice to be entirely independent of its neighbours, but this has the consequence that interpolation cannot be done correctly (sinc interpolation is not correct here). This is immediately apparent when multislice data is reformatted (eg: from original axial to coronal or sagittal): the resulting reformatted views look much less visually pleasing that the reformatted views produced from a 3D volume acquisition with the same nominal z voxel spacing. This problem can be partially addressed by using overlapping slices. These difficulties in interpolating images is worsened by the inconsistency between adjacent slices, such that there is likely to be slight motion between the slices, and motion artefacts are likely to be different between slices.

A further point to consider when interpolating images is that, while images reconstructed using a FT are, by definition, band limited, by the time the data is exported, it has typically been truncated (to remove the greatest wrap-around artefacts), and the image is in modulus format, which distorts the ringing at the dark side of boundaries. This means that correct interpolation is not possible because the data is not band-limited.

Artefacts

Another important limiting factor in many image analysis and visualization problems is artefacts. With increasing use of higher field magnets, and array coils, artefacts are increasingly important. There are three key artefacts that are briefly considered here.

Geometric distortion. This arises due to imperfections in the scanner hardware, and the magnetic properties of the patient. Distortion, however, does not arise equally in all directions. It is greatest in the readout direction in spin-warp imaging (spin echo, gradient echo etc.), but greatest in the blip, or phase encoding direction in echo planar imaging. There are important distinctions between distortion caused by B_0 in homogeneity, and that caused by gradient behaviour linearities. Gradients can introduce linear or non-linear distortion. The linear distortion leads to scaling errors, which can be reduced by scanner calibration. Non-linear distortions lead to position-dependent distortions: if a patient is moved with respect to the isocentre of the magnet, the gradient-induced distortion changes, and this can be a major confound in longitudinal imaging studies unless care is taken to position the patient in the same place relative to the isocentre in each case. This distortion can be mathematically modelled from the gradient design, or quantified using a phantom, and corrected using post-processing. Such post-processing, however, is not always available on scanners, and when available, it may only be 2D, leaving substantial distortion in the 3^{rd} dimension. Heterogenous magnetic properties of the object being imaged can cause B_0 inhomogeneity that leads to distortion that is object dependent. This cannot be corrected by modelling the properties of the magnet, nor by phantom experiments.

Intensity distortion. An identical sample of tissue, in two places in the field of view of an MR image, will have a different intensity. This is a major problem for image analysis, and especially for image segmentation and any techniques (such as visualization) that rely on segmentation. I
Local intensity distortion arises in association with geometric distortion caused by B_0 inhomogeneity, because of the varying magnetic properties in the object. More slowly varying intensity distortion is caused by the interaction of the RF coils with the object. In 3D volume scans, it is also caused by the RF excitation pulse used to excite the volume of interest, but prevent other parts of the body aliasing into the reconstructed image. Intensity distortion is an increasingly important problem to deal with because of the widespread use of higher field magnets, and array coils. At higher field, intensity distortions caused by non-uniformity in the transmit B1 arise, due to wavelength effects that are sometimes called dielectric resonance or field focusing. As a result, birdcage coils will produce less uniform B1 across the object, with the consequence that images from at 3T scanner are noticeably less

uniform than from a 1.5T scanner when using traditional bird-cage coils for transmission. Array coils are becoming widely used in parallel imaging. Their benefits include higher signal to noise ratio, reduced geometric distortion in EPI, and reduced scan time using SENSE or SMASH. These coils are inherently less uniform that birdcage coils, with the consequence that intensity distortion is greater. It is possible to measure the coil sensitivity in situ, and correct the images using techniques such as Clear, pre-scan normalization or Pure, leading to greater intensity uniformity, but the consequence of this is spatially varying noise, which may itself cause problems for image analysis (eg: segmentation algorithms that make assumptions about the noise properties).

Motion artefacts. Motion artefacts can be caused by bulk motion of the subject during scanning, or by physiological motion such as cardiac and respiratory motion and blood flow. Motion causes local inconsistencies in k-space, which result in distributed artefacts called ghosting in the reconstructed image. The appearance of the motion artefact depends not only on its source, but at what time in the scan it happens, and whether it is periodic. In normal Cartesian imaging sequences, the motion artefact is apparent in the phase encoding direction, and for a 3D scan, is typically greatest in the slow phase encoding direction. Motion artefacts can substantially degrade images, and are a major cause of difficulty in image analysis.

These artefacts have important implications for applying geometrical mappings to images. The geometric distortion and motion artefacts are direction dependent, while the intensity distortion is caused by a mixture of patient and instrument factors, which are related to the geometry of the acquisition and patient positioning. When a geometrical mapping is applied to an image, this information about acquisition geometry is lost, which can cause difficulties when using mappings to bring images into a common reference space.

Visualization

It might seem slightly perverse to start a section on MRI visualization by talking about CT, but, CT has often driven developments of rendering techniques and products. Increasing use of multislice CT scanners has lead to increasing clinical (and therefore vendor) interest in rendering techniques. This is driven by the large amount of data that a modern CT scanner can collect, eg: a 64 slice CT scanner can acquire 150+ slices per second, and interpolation during reconstruction can be used to increase this number. Scans encompassing a large field of view (eg: entire spine, or aortic arch to top of head) are becoming common place, and the many hundreds of slices acquired in each scan simply cannot be printed onto film and viewed on light boxes as was done in the past. But CT is much easier to render than MRI. While CT has less good soft tissue contrast, three key structures can be segmented easily: tissue-air interface eg: skin, bronchi, colon (when gas filled); the interface between soft tissue and cortical bone; the interface between blood vessels or cardiac chambers with good contrast filling, and soft tissue. Segmentation in MRI is a much more difficult task, with the exception of contrast angiograms, and the increasing use of array coils and higher field strength magnets seems, at the moment, to be making this segmentation task harder rather than easier due to greater intensity inhomogeneity (array coils without sensitivity profile correction, or dielectric resonance) and spatially varying signal to noise ratio (a problem with array coils, especially if high accelerator factors are used with techniques such as SENSE).

Two distinct approaches to rendering tomographic data are *surface rendering* and *volume rendering*.

Surface rendering

Surface rendering involves pre-segmentation, then rendering a polygonal representation. Surface representations can be manipulated fast on modern graphics cards, giving interactive rendering. They are also easy to use for quantitative purposes (eg: making measurements on the surface such as distance, area or curvature), and they can be texture mapped with the original voxel intensity values – subject to suitable interpolation - if desirable. The de-facto standard technique for generating the surface is the marching cubes algorithm (eg: as implemented in VTK). This technique is designed to extract an isosurface from a volume dataset. It does this using linear-interpolation to give sub-voxel precision in the surface fit. This works well for CT, but not for MRI, because intensity inhomogeneity leads to ambiguous classification, and absence of the high-contract boundaries found in CT makes noise more problematic. As a result, it is common to segment MRI images into binary volumes of interest. This might be done by manual tracing in each slice, by semi-automatic techniques such as region growing and mathematical morphology, or by more automated approaches like statistical classifiers, active shape models, or segmentation propagation. If the marching cube algorithm is run on a binary volume segmented from an MR image, then the interpolation stage cannot work, and the resulting rendered image is very blocky. Two work-arounds are widely used: blurring the binary volume prior to running the marching cubes algorithm, or smoothing the extracted surface. In both cases, detail can be lost, and quantitative measures that might be made on the surface will be distorted.

Volume rendering.

Volume rendering involves direct rendering of the voxels, with on-the-fly classification into structures of interest. A ray is cast from the viewer into the volume. At each sample point along the ray, the contribution of that sample to the screen pixel colour is determined from the voxel intensity, which, by means of a look-up-table, is classified into red, blue, green and opacity values. The opacity can be used to add some uncertainty to the exact boundary intensity value – leading to a more fuzzy rendered image than is normally obtained from surface rendering. Where partial volume effects or image intensity uncertainties are substantial, this fuzziness can be considered a more realistic representation of the underlying data. The volume rendered representation works well for CT, where the classification stage can be done on raw intensities. With MR, there is a need for an intermediate segmentation step, to assign voxels with occupancy values for the various tissues of interest. Then volume rendering can be achieved. But the difficulty in getting good quality segmentations – even fuzzy ones like required here – remains an important obstacle to routine rendering of MRI data. The exception to this is the widely used maximum intensity project techniques, used primarily on angiograms, which simply identifies the brightest voxel along each ray (sometimes with some local averaging in the vicinity of this brightest voxel to reduce sensitivity to outliers).

Non-scalar voxel values.

The discussion of visualization so far has been limited to scalar data. MR images be vector valued (eg: velocity encoded images) or tensor valued (eg: diffusion tensor images).Image analysis algorithms can also generate vector valued images, such as the deformation maps produced by non-rigid registration algorithms. It is also frequently desirable to render this sort of data. The most common approach is to use standard computer graphics techniques such as arrows, streamlines and hyperstreamlines, which are widely used on other scientific data. These approaches are not necessarily optimal – for example they do not properly take account of the properties of the image data. It is likely that this will be the area of visualization that is the focus of research activity in the next few years.

Conclusions

This article has described the properties of MR image data that are important for geometrical mappings, and visualization. Many tools are available for visualization and transformation of images, and users may find it helpful to consider the acquisition related issues discussed here when applying these techniques to their data, and interpreting their results.