Motion and Functional MRI

Informal notes for the Boston'96 Workshop on Functional MRI

Robert W. Cox, PhD

Biophysics Research Institute
Medical College of Wisconsin

Part I: Registration of Rigid Motions

Movement is unavoidable in living human subjects. Since the FMRI (*16*) BOLD effects that we seek are relatively small — 1–5% signal changes — even small motions are important. A simple and ubiquitous case is rigid head motion. Typical T_2^* weighted images show 10–20% brightness changes between adjacent voxels in parenchyma (and 70–80% changes at the edge of the brain). Thus, motions of $\frac{1}{10}$ voxel dimension will produce 1–2% signal changes — on the same order as FMRI signals. For 64×64 images with 200 mm FOV, this means that as little as 300 μm of head motion is an issue. Higher resolution imagery is even more sensitive.

In the present context, *image registration* refers to the problem of aligning the time series of MR images in an attempt to minimize the effects of movement. I will not address the problem of aligning the images from multiple subjects — that is the subject of the "Comparing Brains" session at this meeting. I will use the word *image* interchangeably to mean a 2D or 3D array of voxel intensities. The examples I give will be from 2D image time series acquired with single-shot EPI.

Effects of Registration on the Functional Detection Process

Random motions of the subject will, in any single voxel, look like an additional noise source. If this is the dominant type of movement, registration should reduce the effective noise level, and the number of activations detected should increase.

Stimulus correlated motions of the subject will masquerade as brain function, since the signal changes will be approximately synchronized with the stimulus changes (*5*). If this is the dominant type of movement, registration should reduce the number of activations detected.

Is Motion a Problem in Your Data?

Yes. The only real question is how much of a problem. I strongly recommend viewing the time series of images in 'cine' fashion. The human eye is quite good at seeing $\frac{1}{4}$ voxel motions. Various free software packages can be used for this. One of the fastest animation display utilities for Unix/X11 is *xanim* (*15*).

One common marker for stimulus correlated motion is the presence of large regions of "activation" along the periphery of the brain. A nodding motion of the head is one of the most common types of movement. If correlated with the stimulus, this will produce "activated" regions all along the superior edge of the brain. This is another thing to always look for.

Pairwise Registration Methods

Up to the present, most researchers have concentrated on registering pairs of images. For FMRI purposes, a "typical" image is selected from the time series, and all other images are registered to this. The alternative of attempting to put all the images back together simultaneously has not been well explored, as yet.

Some mathematical notation: the two images are $I(\mathbf{x})$ and $J(\mathbf{x})$. They are presumed to be related by a geometric transformation $\mathbf{T}[\mathbf{x}]$, so that $I(\mathbf{T}[\mathbf{x}]) \approx J(\mathbf{x})$. The problem is to find the transformation $\mathbf{T}$. For 2D rigid body motions:

$$\mathbf{T}[\mathbf{x}] = \begin{bmatrix} \cos\Delta\phi & \sin\Delta\phi \\ -\sin\Delta\phi & \cos\Delta\phi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} .$$

This transformation is parametrized by $(\Delta x, \Delta y, \Delta\phi)$. The analogous transformation in 3D has six parameters: 3 translations and 3 rotations.

Most registration procedures can be characterized by three algorithmic choices:

(a) How to measure the mismatch E (for error) between $I(\mathbf{T}[\mathbf{x}])$ and $J(\mathbf{x})$, so that the algorithm can know when the transformed I is close to the template J?

(b) How to search for the parameters that define $\mathbf{T}$?

(c) How to "reslice" I to the new grid once $\mathbf{T}$ has been determined?

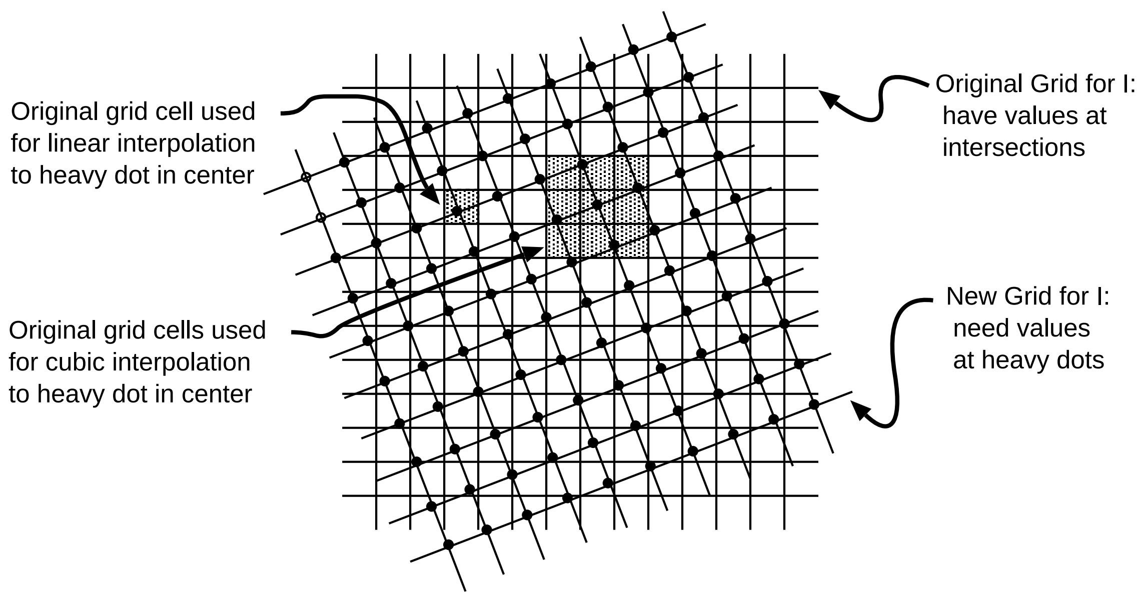

Figure 1. The need for interpolation after finding the motion parameters.

Vast numbers of papers have been written just on these types of registration methods, since each question has at least a few reasonable answers in any given situation. Below, I will describe three such approaches that have been applied to FMRI data.

But First: A Word on Interpolation Methods

As Fig. 1 shows, it is necessary to compute values of $I(\mathbf{x})$ at locations $\mathbf{x}$ that do not correspond to the given data. The simplest way to do this is to find the point on the original mesh that is closest to the desired point, and use the value of I at that original mesh point. This is called *nearest neighbor* interpolation.

The next simplest method is use all the original grid points that immediately surround the desired point. Within that original square (cube in 3D), this technique approximates the values of $I(\mathbf{x})$ by a product of linear functions: $I(x, y) \approx (a + bx)(c + dy)$, where the four parameters $\{a, b, c, d\}$ are chosen to fit the known values at the four enclosing grid points. This method is called *linear* (or multilinear) interpolation.

Higher order interpolation schemes can be used. *Cubic Lagrange* interpolation, for example, uses the 9 grid squares (27 grid cubes in 3D) that surround the desired point. A product of cubic polynomials is fit to the 16 (64 in 3D) known image values at the corners of these squares, and then I can be computed at any interior point. A related method is *cubic spline* interpolation, which computes smoothly fitting cubic polynomials over all the original grid squares at once.

All interpolation methods can be built up from *cardinal basis functions*. The cardinal basis function centered at a given original grid point is the definition of how an image that is 1 at that grid point and 0 at all others is to be interpolated into intermediate points. The final image is then constructed by adding together all the basis functions from each original grid point, each one scaled by the intensity at its original grid point. Figure 2 shows the basis function for 1D linear interpolation. One property that it shares with all cardinal basis functions is that it is zero at all other original grid points except the one at which it is based. In this way, interpolation that falls exactly onto an original grid point will always preserve the original value.

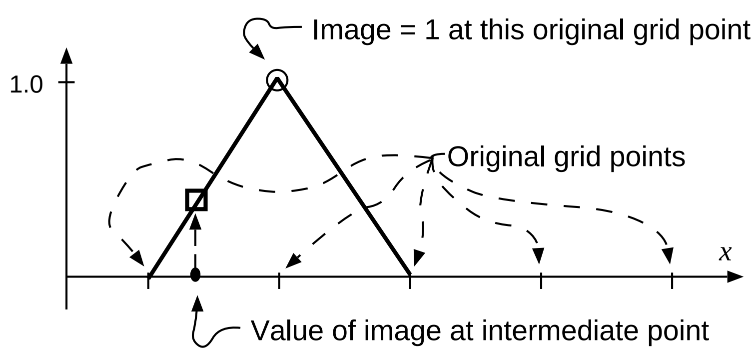

Figure 2. Basis function for linear interpolation.

Sinc interpolation uses the function $\operatorname{sinc}(x - x_{\text{base}}) = \sin(x - x_{\text{base}})/(x - x_{\text{base}})$ as the basis function. If the true image $I(\mathbf{x})$ is bandlimited (*i.e.*, its Fourier transform is zero outside of a finite rectangle of k-space), then sinc interpolation will reconstruct the intermediate values correctly. Since MR image data are acquired directly in k-space, there is no information in a typical $I(\mathbf{x})$ about higher spatial frequencies. Thus, sinc interpolation has some theoretical advantages, since it neither discards low frequency information, nor "makes up" high frequency information. Figure 3 shows the basis function for 1D sinc interpolation, and illustrates how one point influences points far away. This nonlocality makes straightforward implementation of sinc interpolation quite slow, since the output at any given intermediate point depends on the input at many original grid points. There are less obvious methods that can speed the process up.

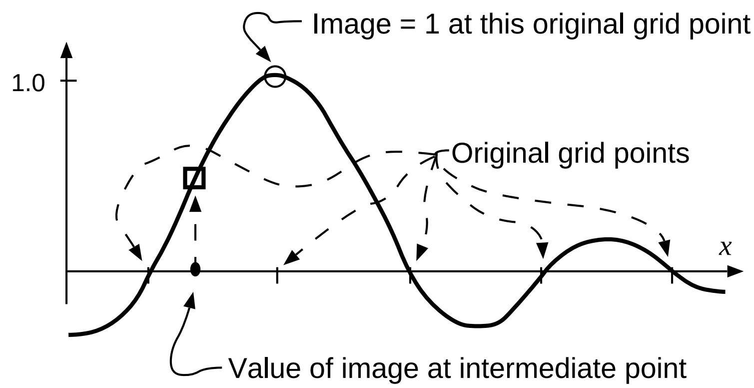

Figure 3. Basis function for sinc interpolation.

Now a Word on Edges

An edge — a very intuitive concept — can be defined as a region in an image where the intensity changes significantly in just a few voxels. In image alignment using image intensities, we expect that edges are the most important locations. As mentioned above, image intensity in parenchyma may change 10% between neighboring voxels, but at the edge of the brain may change 70%. A displacement of 0.1 voxel may then cause a change of 1% in parenchyma and 7% at an edge. If the SNR of the image is 100, then a single 1% change is hard to see (being at the same level as the noise) , but a 7% change is quite significant.

At Last: Some Details About Registration Methods in FMRI

I will only discuss alignment methods that are based on image intensities. An alternative approach that I won't further mention is *feature-based registration.* In these methods, the first step is to use some sort of pattern recognition scheme to select "important" features in both images. The second step is to try to match the locations of these featues by adjusting the motion parameters (Δx, Δy, etc.).

Two of the methods below (1 and 3) are available in widely distributed software packages (*15*). The other one (2) is included here because of the power claimed for it by its creators.

FMRI registration 1: Woods *et al.* (*AIR*) (*13*)

The idea here is to define two images as being effectively the same when their voxel-by-voxel ratio is a constant. If the ratio isn't constant, then by adjusting the motion parameters (*i.e.*, by shifting one image slightly), *AIR* will try to make the ratio more nearly constant. When there is no more improvement — the ratio is as constant as the program can make it — the algorithm stops.

To put the above in mathematical terms, define $r(\mathbf{x}) = I(\mathbf{T}[\mathbf{x}])/J(\mathbf{x})$. Next, compute $\overline{r}$ and σ_r (mean and standard deviation of r) over all voxels $\mathbf{x}$ such that $J(\mathbf{x}) \geq 0.215 \max J$. If $I(\mathbf{T}[\mathbf{x}])$ is truly aligned to $J(\mathbf{x})$, then $r(\mathbf{x})$ should be nearly a constant, and so σ_r should be small. Thus, take the measure of mismatch to be $E = \sigma_r/\overline{r}$.

In the first version of *AIR*, Newton's method was used separately in each parameter of $\mathbf{T}$ in order to minimize E; that is, Δx would be adjusted first (if the E was most sensitive to this parameter), then Δx would be fixed and some other parameter adjusted in its turn. Later versions may use a more elaborate multidimensional minimization scheme — one that varies all the motion parameters at once. Such schemes are generally more efficient, but harder to program.

The first version of *AIR* used linear interpolation between original grid points to get the values on the new grid points. The new version offers sinc interpolation as an option.

FMRI registration 2: Hajnal *et al.* **(*6*)**

Instead of using the voxel-by-voxel ratio of two images to define when they are the same, this method uses the voxel-by-voxel difference between two images. To be precise, the error E is defined as the L^2 mismatch between the images:

$$E = \sum_{\mathbf{x}\in\text{brain}} \left[I(\mathbf{T}[\mathbf{x}]) - J(\mathbf{x})\right]^2 \ .$$

If the only difference between I and J were rigid body movement, then when the proper $\mathbf{T}$ was found, E would be zero. Of course, this won't happen because of noise in the images, but we can expect that varying $\mathbf{T}$ until E can't be made smaller will do a good job of alignment.

The Levenberg-Marquardt algorithm is used to minimize E over the set of parameters of $\mathbf{T}$. This technique is a hybrid of the steepest-descent and Newton methods, and operates on all parameters at once. One advantage of this method is that it can give estimates of the precision with which the parameters of $\mathbf{T}$ have been found. These estimates are based on a statistical model for the noise in the images.

In Ref. (*6*), sinc interpolation is used to reslice I onto the new grid. This point is emphasized in the paper, and the authors clearly consider this to be of prime importance.

FMRI registration 3: Friston *et al.* **(*SPM*) (*3*)**

The difference between this method and the one above is in the details — where the devil is also said to be. The L^2 norm is again used to measure image mismatch. The images are assumed to be "smooth" (if necessary, are lightly blurred to ensure this). Then by linearizing the images (*e.g.*, $J(\mathbf{x} + \boldsymbol{\delta}) \approx J(\mathbf{x}) + \nabla J(\mathbf{x}) \cdot \boldsymbol{\delta}$) and assuming the motions are small, it is possible to set up a set of linear equations to solve for the unknown parameters of $\mathbf{T}$. This method is approximately equivalent to the first step of a steepest descent minimization of E, and so is similar to the first step of the Levenberg-Marquardt algorithm used in the previous method.

If desired, the procedure can be iterated after the images are resliced (*9*); however, this does not seem to be advocated in Ref. (*3*). Linear interpolation seems to be used for resampling to the new grid. The combination of single-step linearization and linear interpolation makes this method relatively simple to implement and very fast to execute.

Advanced Issues in Registration: Interpolation

Interpolation of an MR image to a new grid is really a reconstruction issue. The ideal operation would be to determine the movement parameters and then reconstruct the original raw data directly onto the desired new grid. For k-space spiral imaging, this is no more difficult than the original reconstruction, which is carried out by "gridding" the spiral data onto the reciprocal k-space grid of the desired x-space grid, and then using the FFT.

For rectangular scan k-space imaging, a similar procedure could be carried out by interpolating the data to the desired reciprocal grid (which will be rotated from the acquisition k-space grid). This is widely — but incorrectly — believed to be necessary for efficiency since the FFT algorithm is limited to transforming an array of data from one regular grid to its reciprocal grid. In fact, a generalization of the FFT called the *chirp z-transform* will efficiently transform an array of data from one regular grid to any other regular grid (*12*). This is equivalent to a periodic version of sinc interpolation, using the function $\sin(\pi x/h)/N \sin(\pi x/Nh)$ as the basis function (h = grid spacing, N = number of grid points).

Advanced Issues in Registration: 3D Imaging

The dominant form of acquisition in FMRI is currently 2D multislice imaging. There are two problems with this in regards to volumetric (3D) registration. The first is that the slices are not gathered at the same time. A typical T_R for whole brain acquisition is 5 s. Due to the interleaving of slices, this means that adjacent slices are acquired 2.5 s apart. Assembling all these slices together into one 3D volume, and registering them to the next (5 s later) volume, is artificial — it is pretending that there was no significant motion in the 2.5 s interval, but that there might have been in the 5 s interval:

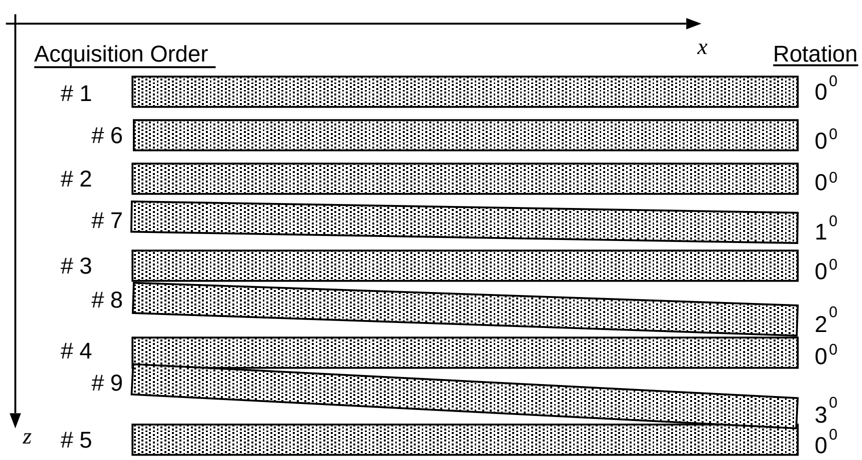

Figure 4. Multislice acquisition with head rotation starting at the seventh slice acquisition.

What is really needed for this type of acquisition is a "slice-into-volume" type of registration.

Another problem with 2D multislice imaging is interpolation in the z direction. As I said before, interpolation is a form of reconstruction. For 2D imaging, reconstruction in xy is via the Fourier transform, and reconstruction in z is via the slice selection pulse. Rotating and shifting in the xy-plane is just a matter of using the Fourier transform properly. Proper interpolation in the z direction requires knowledge of the slice selection profile.

For these (and other) reasons, the future of whole brain FMRI may lie with true 3D acquisitions. This is not likely to be done using single shot techniques, due to the lack of achievable resolution within the T_2^* signal decay time. With multishot techniques, the issue of head movement between shots still arises. Uncorrected, these will lead to image blurring and other artifacts. Compensation for such motions will require the gathering of redundant information with each shot (*e.g.*, navigator echoes) that can be used to assemble the multishot information correctly.

Advanced Issues in Registration: Limits of Accuracy

How well, in principal, can two MR brain images be registered by rigid body motions? There are (at least) two answers to this question. The first answer comes from treating the issue as a statistical question: how much information about the motion parameters is in the noisy images? The second answer comes from treating it as a practical question: how much does the brain behave like a rigid body?

One way to answer the statistical question is to formulate a mathematical model for the images to be registered, and then apply the Cramér-Rao theorem, which establishes bounds on the accuracy of parameter estimation based on noisy data. Some simple model assumptions:

$$\begin{aligned}
J(\mathbf{x}) &= F(\mathbf{x}) + \eta_1(\mathbf{x}) \\
I(\mathbf{x}) &= F(\mathbf{S}[\mathbf{x}]) + \eta_2(\mathbf{x}) \qquad (\mathbf{S} = \mathbf{T}^{-1}) \\
\eta_i &= \text{zero mean, Gaussian, with autocovariance function } \Sigma(\mathbf{x}, \mathbf{x}') \\
\Sigma(\mathbf{x}, \mathbf{x}') &= \mathbf{B} * \Sigma_0(\mathbf{x})\delta(\mathbf{x} - \mathbf{x}') * \mathbf{B} \\
\Sigma_0(\mathbf{x}) &= \text{noise variance as function of position} \\
\mathbf{B} &= \text{blurring operator (to approximate intervoxel noise correlation)} .
\end{aligned}$$

With this model, one can estimate $\Sigma_0(\mathbf{x})$ from data, and then compute the Cramér-Rao bound on estimation of motion parameters. For some sample 64×64 EPI data at 1.5 T, the results come out to be that the images can be registered to within 25 μm ($\pm 1\,\sigma$), *assuming the mathematical model is correct.* Interestingly, this result is more than double the value that comes from repeating the same calculation using the assumption of spatially uniform variance (*i.e.*, that $\Sigma_0(\mathbf{x})$ is a constant). This indicates that testing a registration algorithm with such simple noise may be misleading.

For a model such as the one above, the Cramér-Rao bound can asymptotically be achieved by maximum likelihood estimation. This tells us (roughly) that each datum should be weighted inversely proportional to its variance. Figure 5 shows how much the voxel-wise standard deviation of an image time series varies between locations.

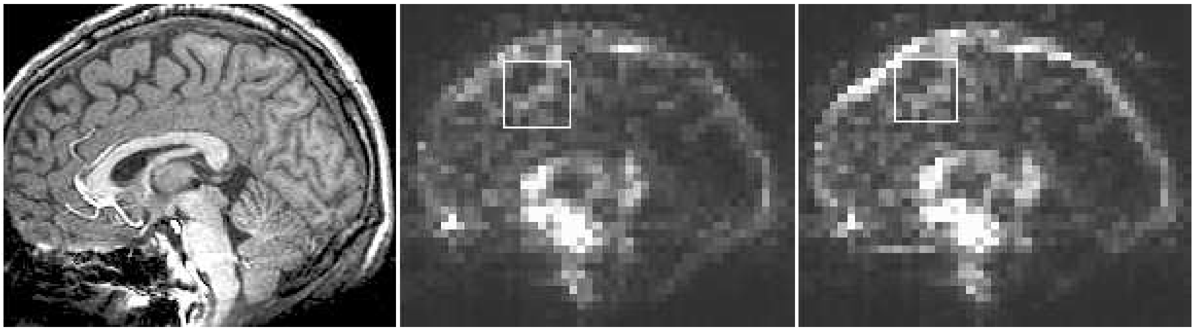

Figure 5. Anatomy; standard deviation with (center) and without (right) registration.

Much of the variance in the unregistered time series (right image) is at the edge of the brain, and is removed by simple pairwise registration (center image). This suggests a two step strategy for registration: first, align the image series in some fashion, then compute the voxel-wise variance of the resulting time series and realign, weighting each voxel by the inverse of its variance. This procedure may help avoid the pitfall of trying to register voxels that are activated. For example, the brighter voxels in the center of the boxed region correspond to functionally activated voxels. Their relatively high variance means that they wouldn't be counted much in the second registration pass.

The practical question remains: how good is the model of rigid body motion? One way to approach this question is to find out how much the brain deforms even when there is no overall motion of the head. This has been investigated in recent years using rapid imaging sequences (*10, 11*). The brainstem moves about 500 μm with the cardiac cycle, and can also move up to 2000 μm with forced inspiration or breathholding. At the other end, motion of the cerebral cortex with the cardiac cycle is only 10–20% of the brainstem motion. Thus, the brain's shape changes during the cardiac cycle. This distortion represents the practical limit to image time series alignment via rigid translations and rotations — my guess is that this limit is on the order of 100 μm in the cortex and 200+ μm in the deep gray matter.

One approach to minimizing the brain distortion problem is to gate (or double gate) the MR acquisition to always occur at the same phase of the cardiac cycle. In the hands of the MGH group, this has reduced the image variance in the brainstem region and made possible the detection of functional activation in this important part of the CNS.

To the extent that the brain distortions repeat themselves with each heartbeat and breath, then some of their effects may be removed by filtering operations (after registration for rigid motions). See Part III for outlines of two such approaches.

Advanced Issues in Registration: Where is the Information?

That is, what parts of the image contain the most information about how to reposition it? There are two intuitive guidelines:

(a) The edges should be important, since moving an edge will cause a big change in nearby voxels.

(b) Regions with large variance are the least important, since their values will change "randomly" the most from image to image.

But what about edge regions with large variance? An extension of the mathematical model given in the previous section allows the information content (relative to the motion parameters) of any given voxel in the whole ensemble to be estimated using information theory and the Cramér-Rao bound.

Figure 6 shows the results of such a calculation, again applied to a 2D 64×64 EPI time series acquired at 1.5 T. The edge image is bright where we expect information about brain position to be high — guideline (a). The standard deviation image is bright where we expect information about brain position to be low — guideline (b). The information image is the result of a formal calculation. It looks like the edge image, except it is faded out where the standard deviation image is bright. This is the quantitative interaction of the guidelines above. In the future, such information maps may be useful for the design of registration algorithms.

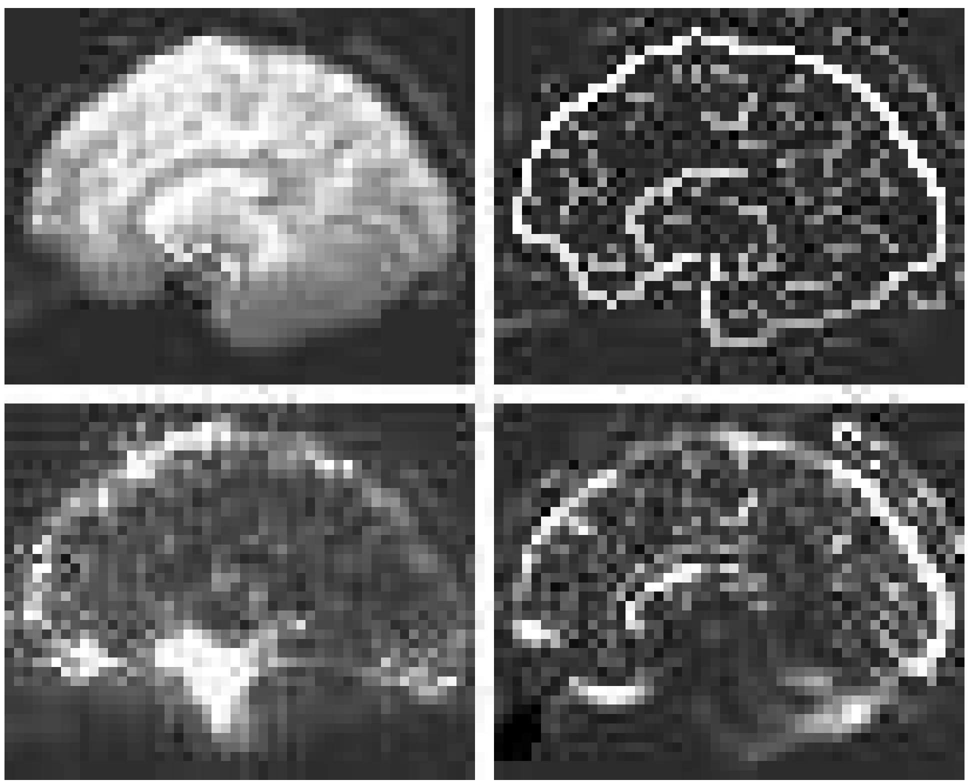

Figure 6. Upper left: One echo planar image from time series.
Upper right: Sobel edge detection applied to echo planar image.
Lower left: Standard deviation from time series.
Lower right: Information about rigid motion parameters.

Part II: Other Image Changes Due to Rigid Motions

Due to the nature of the MR imaging process, rigid motion of the object being scanned does not just imply corresponding rigid motion of the image. A crude analogy: the MRI process is like photographing an object through a wavy distorting piece of glass. If the object moves behind the glass, the distortions are different because the light passes through a different set of ripples in the glass. Even if the distorted image is then moved back to where it "ought to have been", the different distortions remain, and the registered image is not identical to what would have been obtained if there had been no movement.

Imaging Physics in 30 Seconds (or so)

Magnetic resonance images are formed by the interaction of two magnetic fields with water protons. These fields are the 'static' B_z field, which makes the protons resonant with the radiofrequency B_1 field. B_z can in turn be thought of as comprising two parts. The first, B_0, is the large field produced by the superconducting magnet. The second, $\mathbf{G}(t) \cdot \mathbf{x}$, is the time-dependent gradient field that is used to encode position information into phases and frequencies. The B_1 field is used to excite the water protons' spins. The radiation emitted by the spins as they relax from the excitation and precess at frequency $\gamma B_z(\mathbf{x}, t)$ is the raw data used to compute the image. The readout can be accomplished by the same RF coil that was used for excitation, or by a separate coil.

Nonuniformities in any of these scanner generated fields will cause signal changes in the MR signal when a subject's head moves within the scanner, and thus moves the nonuniformities from one anatomical location to another. In principal, these fields can be mapped and their effects compensated for during reconstruction. If the subject's movements are small enough, then estimates of the motion parameters using methods discussed in Part I could be used to guide the elimination of these distortions using the previously mapped scanner fields. In practice, this will be complicated by the fact that other effects, discussed below, also cause movement-dependent changes in the image.

Echo-planar imaging is particularly sensitive to distortions in B_z. This is due to the long duration of the readout from a single excitation with this pulse sequence. Since the nuclear resonant frequency is proportional to B_z, errors in frequency have longer to accumulate into errors in phase with EPI. In rectangular scan EPI, the effects of nonuniform B_z are seen as distortions of the image in the phase-encoding direction and as variable intensity ghost images. These effects arise from the regular nature of the back-and-forth progression in k-space with this type of EPI. In spiral scan EPI, this regularity is lost, and nonuniform B_z results principally in nonuniform blurring of the images.

Susceptibility Induced Changes to B_z

Even if there were no imperfections in the scanner system itself, nonuniform B_z would still be an unavoidable fact of life. This is due to the fact that tissue has nonzero magnetic susceptibility. This means that if it is placed in an external magnetic field B_0, the total magnetic field will be the sum of B_0 and an extra field produced by the action of B_0 on the electrons in the tissue. What is more, the effect of susceptibility is not strictly local: susceptibility in one place causes changes in B_z at other place.

The equation for the magnetostatic potential is $\nabla \cdot [\mu(\mathbf{x})\nabla\Phi_m] = 0$ (the applied mathematician's union requires me to include at least one partial differential equation in every document I write). In practice, $\mu(\mathbf{x}) = 1 + \delta\mu_1(\mathbf{x})$, with $\delta \approx 10^{-6}$, and with $\delta\mu_1$ being the tissue magnetic susceptibility. Since δ is so small, we expand Φ_m in a power series in δ to find

$$B_z(\mathbf{x}) = B_0 + \delta B_0 \left[\mu_1(\mathbf{x}) + \frac{1}{4\pi} \iiint \frac{\partial\mu_1(\mathbf{x}')}{\partial z'} \frac{z - z'}{|\mathbf{x} - \mathbf{x}'|^3} \, d^3x' \right] + \cdots .$$

This shows that not only does $\mu_1(\mathbf{x})$ directly affect B_z at $\mathbf{x}$, but that $\partial\mu_1/\partial z$ also affects B_z nonlocally. The z-axis here is defined by the main magnetic field. Thus, if the head rotates around x or y (nodding or side-to-side motion), then $\partial\mu_1/\partial z$ will be the derivative along a new direction in the tissue. In turn, the image will be distorted or blurred differently than if no motion occurred.

This effect is somewhat different than the changes caused by head motion through an inhomogeneous B_z generated by the scanner. In the present case, translation of the head causes no effect, but rotation causes the actual B_z field to change. This cannot be mapped ahead of time, and so will be difficult to correct for, even in principal.

Slice Selection and Spin History

With 2D multislice imaging, out of plane motion will cause a slightly different slice to be excited at each repetition. For example, spins near the edge of a slice may move into a neighboring slice's volume by the time the neighbor is to be excited. If the subject hadn't moved, the stationary spins would be excited once per T_R. With this type of movement, the spins that moved into the neighbor will get excited again after just $\frac{1}{2}T_R$. If T_R is shorter than $4T_1$ (say), the amount of longitudinal magnetization recovery will be different for these moved spins than for the spins in the bulk of the neighboring slice, which will have last been excited a full T_R in the past. This will cause the signal intensity in the neighbor slice to be lower than it would be without motion, since some of the spins being imaged will not have undergone as much longitudinal relaxation. Even after perfect spatial registration, these signal intensity changes will remain.

The obvious way to eliminate this problem is to use true 3D acquisition. Until this is routinely practicable, Friston *et al.* (*4*) have proposed to reduce these intensity changes by linear filtering — essentially, by regressing the data against the movement time series derived from the registration process.

Part III: Physiological Motions

Much of the apparent noise in an FMRI time series arises in fact from respiration and cardiac pulsations. These are true MR signals, just not the signals of primary interest for brain mapping. (One definition of noise: a signal that you don't want for the purposes of your current study.) Since these sources of voxel intensity change occur at rates that are quite different from the stimulus-induced changes being mapped with FMRI, they can be separated by appropriate signal processing.

Reduction of Physiological Fluctuations: Linear Filtering

The most straightforward technique is to observe that the respiratory and cardiac cycles are approximately periodic. In the Fourier domain, this means that their effects are limited to certain frequency bands. If the image time series is acquired rapidly enough to resolve these frequency bands separately from the stimulus-induced signal changes, then linear digital filtering can remove the unwanted fluctuations (*2*).

Linear filtering is equivalent to projection onto the subspace orthogonal to the rejected frequency bands. The widely used correlation method for activation detection is equivalent to projection onto the subspace spanned by the reference vector(s). Thus, this type of linear rejection of certain noise components is easily incorporated into the correlation method simply by making sure that the reference vectors are orthogonal to the undesired subspace.

One difficulty with the linear filtering approach is the high frequency of the cardiac cycle — 1.1 ± 0.1 Hz in typical healthy young subjects. To fully resolve this requires sampling the volume of interest every 416 ms [$= 1/(2 \cdot 1.2\text{Hz})$]. This is feasible with EPI for a few slices, but not for whole brain imaging. Alternatively, one can accept aliasing and sample at a lower rate. As long as the cardiac band is not aliased to overlap with the stimulus band, the filtering can still be carried. This technique requires measuring the frequencies influenced by the cardiac cycle in the individual subject, and then making sure that the aliasing is acceptable. For the purposes of individual studies, such an approach is feasible. For the purposes of group studies, where results from identical runs on many subjects must be compared, such an approach probably cannot work. It would be necessary to adjust the stimulus or data acquisition rate to avoid aliasing overlaps in some subjects, and then these acquisitions would no longer be identical.

Reduction of Physiological Fluctuations: Monitoring Activity

A second approach to the reduction of the effects of physiological involves monitoring the cardiac and respiratory cycles. This may be done by external sensors (*7*) or by using appropriate MR echoes within the data acquisition process (*8*). In this technique, the physiological fluctuations are not explicitly assumed to be narrowband. Instead, each cycle (cardiac or respiratory) is assumed to be a copy of a basic cycle, stretched or contracted in time to match the duration of the individual event (heartbeat or breath — breath cycles are also allowed to scale in magnitude). The effects of each event on the MRI signal intensity in a given voxel are assumed to be the same as in the basic cycle, after allowing for the time stretching or contracting needed to fit the individual event to the basic cycle. Thus, each data point in time is assigned a cardiac and a respiratory phase, as determined by the monitored physiological activity. Then components of the data which are periodic with these phases are removed by a linear least squares fit to a low-order Fourier series. (Note: these fits are referred to as nonlinear in Ref. (*7*), but the formula given therein — Eq. [2] — is linear in the fit parameters.)

This technique requires monitoring of the cardiac and respiratory cycles at a high enough rate to accurately know the physiological phases of every MR image point. With external sensors, this is relatively simple, but requires that the external data acquisition devices be accurately synchronized with the MRI acquisition system. With the use of internal echoes for estimation of the physiological activity, this problem is replaced by the problem of repeating the identical internal echoes fast enough. For single-slice imaging, this is not a difficulty, but for 2D multislice imaging this makes the requisite pulse sequences more complex. Perhaps this is another argument for true 3D acquisitions in FMRI.

Complex Subject Motions

Motion of objects strictly outside of the FOV can still affect the images. In particular, jaw motion and swallowing have been observed to cause artifacts in EPI. These effects are due to the nonlocal impact of susceptibility. Suppose that the nuclear precession frequency field γB_z changes by an amount $\Delta\omega(\mathbf{x})$. Then the signal gathered may be modeled as

$$S(\mathbf{k}) = \int M_\perp(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x} - i\Delta\omega(\mathbf{x})t(\mathbf{k})}\, d\mathbf{k}$$

where $t(\mathbf{k})$ is the time at which k-space point $\mathbf{k}$ is sampled. The distortion of phase — which is ideally just $\mathbf{k}\cdot\mathbf{x}$ — causes changes in the reconstructed image. Work with phantoms at MCW has shown that it is possible to detect these motions by monitoring the phase of the reconstructed complex-valued image (*1*).

Some correction of these artifactual changes can be effected by adjusting the phase of the data to match a base image. Such corrections require a very stable MRI system, so that the image phase remains unchanged between shots if nothing changes in the FOV. A difficulty with this approach in human trials has been that extra-FOV motion is always accompanied by intra-FOV motion. Thus the inter-image phase matching must be carried out in tandem with the inter-image registration.

Part IV: Possibly Controversial Points

For such a non-squishy scientific area, there is a surprising amount of controversy. Besides whatever issues I have raised above, here are a few more items to get the discussion rolling:

1. *How much motion can we tolerate?*

The rigid motions that we seek to correct for are not very large in FMRI — a few voxels, at most. Motions larger than this probably corrupt the signals so much due to other causes (see Part II) that restoration down to the level of functional activation changes is highly problematical. I would be very wary of basing scientific conclusions on or making clinical decisions with such highly massaged data sets.

Very desirable would be some systematic brain mapping studies to determine if such inferences can be reliable in the presence of large movements.

2. *What should a registration algorithm match, and how?*

For the correction of small-ish motions, I *do not* feel that the choice of objective function (E) or minimization algorithm is likely to make much difference. Certainly no one has carried out any extensive study using different registration methods to determine if the brain mapping results will be different if these two algorithmic choices are altered.

3. *What interpolation method is adequate? Best?*

I *do* feel that the choice of resampling method is important. Linear interpolation is very easy to implement, but it is not accurate enough — it tends to blur the final images too much. Cubic and sinc interpolation are far superior in this regard.

Figure 7. Original (left) and repeatedly rotated images.

The image to the left is the original. It was rotated in 10° increments 36 times, back to the original orientation. Cubic interpolation was used for the middle image at each step; linear interpolation for the right image. Even after printing and duplication, the gross blurring from repeated linear interpolation should be clear. The final interpolation step should be taken with a high-accuracy method, directly from the original grid.

4. During an iterative registration process, it is necessary to transform an image using the currently estimated motion parameters. The mismatch between this temporary transformed image and the reference image is then used to update the motion parameter estimates. The original image is then re-transformed, etc. (Never transform a transformed image! Figure 7 should convince you of that.) Therefore:

How accurately does the intermediate transformed image need to be interpolated?

My opinion: linear interpolation is adequate for this purpose. This is based both on practical trials and on computing the Cramér-Rao bound, where the effects of inaccurate interpolation can be modeled by a broader blurring operator $\mathbf{B}$.

5. *Is it adequate to linearize the effects of motion?*

Some workers have proposed simply to stop the iterative registration process after the first step (*4*). This is acceptable for very small movements, but begins to lose accuracy when the displacements are larger than $\frac{1}{2}$ voxel dimension. The graphs in Fig. 8 plots Δx (left) and $\Delta\phi$ (right) obtained by iterative 2D registration (abscissa) vs. Δx and $\Delta\phi$ obtained by single-step registration (ordinate). The original images are from a time series of 64×64 whole brain EPI with 240 mm FOV and $T_R = 5$ s.

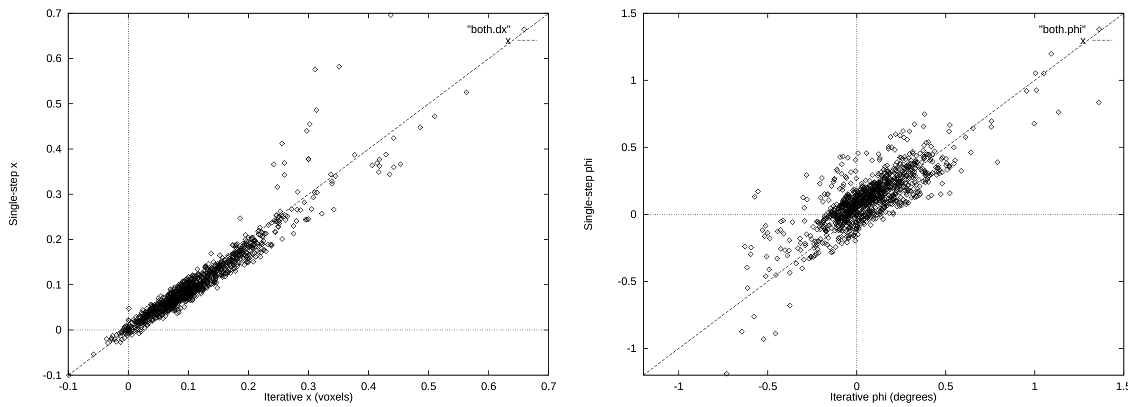

Figure 8. Iterative vs. noniterative estimates of Δx and $\Delta\phi$.

These results are from real motion not monitored by any external devices; which is more correct is *à priori* unknown. Tests with simulated motion — moving an image, corrupting it with noise, then registering it to the original — show similar results.

I believe that iteration is necessary — registration is a nonlinear process. The widespread belief that iteration must be very slow is not correct. The above registrations took 50 ms per 2D image with iteration, and 16 ms in single-step mode (running on a 150 MHz Pentium). Careful implementation of the interpolation and error minimization routines is the most important factor when it comes to speed.

6. *Is it necessary to allow for overall image brightness changes from shot to shot?*

These can be fit by introducing an extra scaling parameter β, so that we say $I(\mathbf{T}[\mathbf{x}]) \approx \beta J(\mathbf{x})$. Then we solve for β along with the parameters in $\mathbf{T}$. An alternative approach is given by the *AIR* package, which uses a goodness-of-fit measurement that is insensitive to β.

7. *How can image realignment software be validated?*

It is not adequate to take an image, move it in software to some other location, add some Gaussian white noise, and then see how accurately the software estimates the motion. Real noise is not white or stationary (*cf.* Figs. 5 and 6), and probably not extremely Gaussian.

8. It is possible to model the geometric transformation **T** as something more complicated than a rigid body motion — **T** could also include extra parameters to allow for warping effects (see Ref. (*3*), for example).

Is it useful to insert some arbitrary distortion parameters into the "registration" process?

The main theoretical difficulty is that realistic distortions (*e.g.*, due to the B_z effects discussed before) will require a lot of parameters to model well. But is there enough information in the data to extract these parameters accurately?

9. *Does it make sense to regress the data (in some sense), after registration, against the movement time series?*

The justification for such a procedure is basically that the effects of Part II may be approximated by a model that is linear in the motion (*4*). The parameters of this model can be fit via regression, and the effects then "removed".

Acknowledgements

I thank Bharat Biswal and Roger Woods for thoughtful and provoking comments.

References

1. Birn RM, Yetkin FZ, and Hyde JS. Artifact in fMRI Caused by Motion Outside the FOV. *Proceedings Fourth Ann. Mtg. ISMRM*, 1770, New York (1996).

2. Biswal B, DeYoe EA, and Hyde JS. Reduction of physiological fluctuations in fMRI using digital filters. *Magn. Reson. Med.* **35**:107–113 (1996).

3. Friston KJ, Ashburner J, Frith CD, Poline J-B, Heather JD, and Frackowiak RSJ. Spatial registration and normalization of images. *Human Brain Mapping*, **3**:165–189 (1995).

4. Friston KJ, Williams S, Howard R, Frackowiak RSJ, and Turner R. Movement-related effects in fMRI time-series. *Magn. Reson. Med.* **35**:346–356 (1996).

5. Hajnal JV, Myers R, Oatridge A, Schwieso JE, Young IR, and Bydder GM. Artifacts due to stimulus correlated motion in functional imaging of the brain. *Magn. Reson. Med.* **31**:283–291 (1994).

6. Hajnal JV, Saeed N, Soar EJ, Oatridge A, Young IR, and Bydder GM. A registration and interpolation procedure for subvoxel matching of serially acquired MR images. *J. Comput. Assist. Tomog.* **19**:289–296 (1995).

7. Hu X, Le TH, Parrish T, and Erhard P. Retrospective estimation and correction of physiological fluctuation in Functional MRI. *Magn. Reson. Med.* **34**:201–212 (1995).

8. Le TH, and Hu X. Retrospective estimation and correction of physiological artifacts in fMRI by direct extraction of physiological activity from MR data. *Magn. Reson. Med.* **35**:290–298 (1996).

9. Lucas BD, and Kanade T. An iterative image registration technique with an application to stereo vision. *International Joint Conference on Artificial Intelligence*, 674–679, Vancouver (1981).

10. Maier SE, Hardy CJ, and Jolesz FA. Brain and cerebrospinal fluid motion: real-time quantification with M-mode MR imaging. *Radiology* **193**:477–483 (1994).

11. Poncelet BP, Wedeen VJ, Weisskoff RM, and Cohen, MS. Brain parenchyma motion: measurement with cine echo-planar MR imaging. *Radiology* **185**:645–651 (1992).

12. Rabiner LR, Schafer RW, and Rader CM. The chirp z-transform algorithm and its application. *Bell System Tech. J.* **48**:1249–1292 (1969).

13. Woods RP, Cherry SR, and Mazziotta JC. Rapid automated algorithm for aligning and reslicing PET images. *J. Comput. Assist. Tomog.* **16**:620–633 (1992).

14. Woods RP, Mazziotta JC, and Cherry SR. MRI–PET registration with automated algorithm. *J. Comput. Assist. Tomog.* **17**:536–546 (1993).

15. Software mentioned herein:

    Xanim is at `http://www.portal.com/~podlipec/`.
    AIR is at `http://bishopw.loni.ucla.edu/AIR/index.html`.
    SPM is at `http://www.fil.ion.bpmf.ac.uk/spm/`.
    My software is at `http://www.biophysics.mcw.edu/`.
    (I may also be reached by e-mail at `rwcox@mcw.edu`.)

    Like most WWW addresses, these are valid as of today (June 6, 1996) only.

16. **FMRI** or **fMRI**? The standard practice in American English is to capitalize acronyms. In my humble opinion, "fMRI" is silly.