Combining Magnetic Resonance Image Planes in the Fourier Domain for Improved Spatial Resolution

Mark A. Hasegawa-Johnson*
mhj@icsl.ucla.edu

University of California at Los Angeles

Abstract

This paper describes a method for combining Fourier transform data from multiple orthogonal image planes in order to create a uniformly-sampled, high-resolution three-dimensional image. A discussion of the problem is followed by a theoretical presentation of the solution, and of its limitations. Results are presented from validation experiments, in which coronal and axial images of the mouth and throat are combined in order to create a three-dimensional image with better resolution than either original image.

1 Problem Statement

The research described in this paper grew out of an ongoing speech production research program at UCLA [3]. In this research program, paid experimental subjects are asked to sustain particular speech sounds while they are imaged in a standard clinical magnetic resonance (MR) scanner. The first goal of the research is to reconstruct the three-dimensional shape of the mouth and throat during each speech sound using information from the MR images.

Reconstruction of the three-dimensional shape of an imaged object would be simplest if the object was imaged on a uniform Cartesian grid. For example, figure 1 shows a grid in which tissue density is measured in cubic voxels which measure 2mm on each side; the subdivided cube shown in the figure represents a block of 27 voxels. The grid shown in figure 1 can easily be constructed using a standard clinical MR scanner by gathering a stack of image slices, each with a slice thickness of 2mm, and each with an in-plane pixel size of 2×2mm.

Figure 2 shows an axial image of the mouth

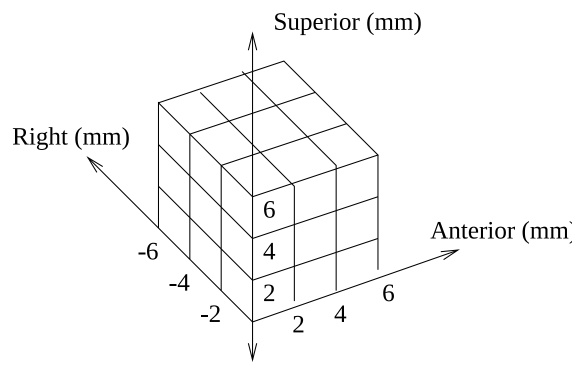

Figure 1: Three-dimensional reconstruction is simplest if the target is imaged using a uniform Cartesian grid. For example, tissue density might be measured in cubic voxels which are 2mm on a side, as shown here.

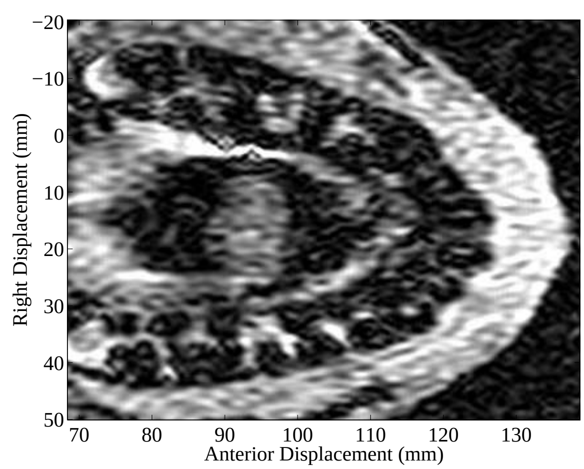

Figure 2: Axial MR image of the mouth, measured using a 2mm slice thickness with a pixel size of 0.9375×1.875mm.

*Address from September, 1999: University of Illinois at Urbana-Champaign.

measured in a standard clinical GE SIGNA scanner, with a 2mm slice thickness, and with an in-plane pixel size of approximately 1×2mm. In this image, teeth and air are black, while the tongue, cheeks, dental gingiva, and other soft tissue are gray or white. The black and white speckles throughout this image are caused by thermal noise. With this much thermal noise in the image, it is impossible to confidently and accurately trace the outlines of the visible soft tissue.

Thermal noise in an MR image is caused by random fluctuations in the magnetic dipoles of all atoms in the bay of the scanner; thermal noise is not dependent on the size of the imaging voxel [5]. Since these fluctuations add incoherently, the RMS strength of the imaging noise N is proportional to the square root of imaging time T:

$$N \propto \sqrt{T} \tag{1}$$

The MR signal is obtained by causing the dipoles in a specified imaging plane to precess with known frequency and phase, and then measuring the signal they radiate. RMS signal strength S is proportional to the number of atoms per image voxel, and therefore to the volume V of the voxel; since signal adds coherently, signal strength is also proportional to imaging time:

$$S \propto VT \tag{2}$$

$$20 \log_{10}(S/N) \propto 20 \log_{10}(V\sqrt{T}) \tag{3}$$

As shown in equation 3, it is much easier to increase SNR by increasing V than by increasing T. For example, an extra 9dB of SNR can be obtained by scaling V by a factor of three, or by scaling T by a factor of nine.

The standard clinical strategy for increasing SNR is shown in figure 3. In this strategy, pixel sizes in the plane of the image are maintained, while the thickness of the plane is increased; in figure 3, the 2mm slice thickness of figure 1 has been increased to 6mm, with the result that the size of each individual voxel is increased from 8mm^3 ($2 \times 2 \times 2$mm) to 24mm^3 ($2 \times 2 \times 6$mm). In the normal jargon of medical imaging, the 3×3 square shown in figure 3 is a "projection" in the z direction of the $3 \times 3 \times 3$ grid shown in figure 1:

$$p_z(x, y) = \sum_{z=1}^{3} a(x, y, z) \tag{4}$$

where x, y, and z are discrete voxel indices in the left-right, anterior-posterior, and superior-inferior directions, respectively, $a(x, y, z)$ is the original uniformly-sampled volume image, and $p_z(x, y)$ is the axial image "projection."

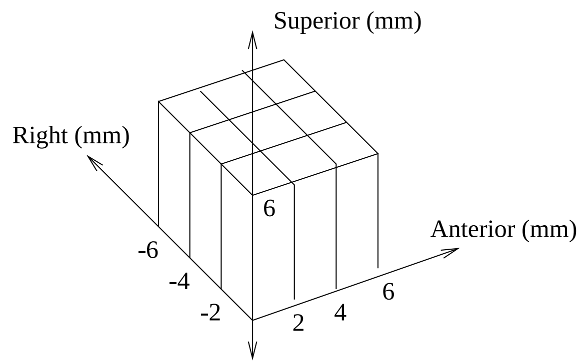

Figure 3: A strategy for imaging the same volume as that shown in figure 1, but with three times the volume per voxel. Pixels in the axial image plane are still 2mm squares, but the slice thickness perpendicular to the image plane is increased to 6mm.

Long thin voxels of the sort shown in figure 3 have a number of practical advantages. Radiologists like them, because they give a higher-resolution two-dimensional image slice. A small in-plane pixel allows certain computational advantages, such as the use of large FFT sizes to compute the image. Finally, all else being equal, a thicker image slice can be selected using a shorter RF pulse, and/or with a weaker static field gradient. For all of these reasons, it is very difficult to obtain MR images from a clinical scanner using any shape other than the long, thin shape shown in figure 3.

In order to reconstruct three-dimensional objects using MR image data, therefore, it is necessary to compensate in some way for the long, thin shape of the image voxels produced by a standard clinical scanner.

2 Proposed Solution

Figure 4 is a screen shot from a viewing, editing, and reconstruction package called "mrcat" or "CTMRedit," which was developed as part of ongoing speech production/biomedical imaging research at UCLA [1, 2][1]. The user of this package edits the outline of up to six different-colored regions of interest in the main image (the large axial image). Yellow "locator lines" (not visible in the

[1] Available for download from http://www.icsl.ucla.edu/~spapl/CTMRedit/index.html.

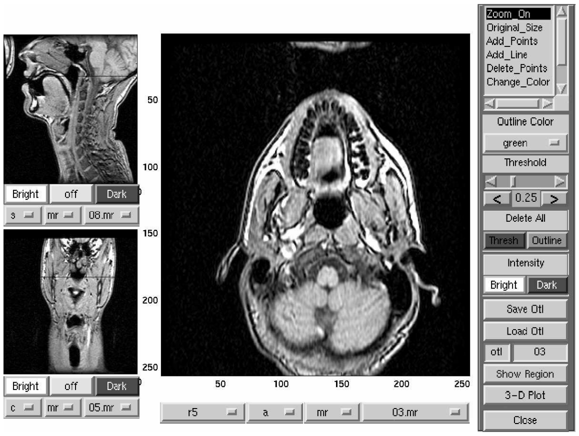

Figure 4: Region of interest editing tool "mrcat," showing sagittal, axial, and coronal MR images measured with a 5mm slice thickness, and an in-plane pixel size of 0.9375×1.875mm

figure) show the location of the main image plane in each of two small locator images (sagittal and coronal images in the figure); the user can take advantage of information in the locator images to better identify structures visible in the main image.

CTMRedit is easiest to use if high-resolution images exist in all three orthogonal image planes. For this reason, all subjects imaged in our current research are imaged in at least two orthogonal planes, and many subjects are imaged in all three orthogonal planes. For example, figure 4 shows sagittal, coronal, and axial images of an experimental subject producing the sound $/\epsilon^r/$ as in the word "bird." A total of 16 sagittal, 23 coronal, and 23 axial images were taken of the subject performing this task, with a 5mm slice thickness, and no gap between successive images in each stack. Images were collected in several breath-holds over a period of about ten minutes; the subject was not told when the imaging plane changed from sagittal to coronal, or from coronal to axial.

Recall from equation 4 that the sagittal, coronal, and axial images shown in figure 4 can each be viewed as the projection of an underlying, uniformly sampled image $A(x, y, z)$. It is possible to estimate $a(x, y, z)$ by combining information from the three projections using the projection-slice theorem [4]. If $A(k_x, k_y, k_z)$ is the three-dimensional Fourier transform of the uniformly sampled image $a(x, y, z)$, and $P(k_y, k_z)$ is the two-dimensional transform of a projection $p_x(y, z)$ created by summing $a(x, y, z)$ in the x direction, then the projection-slice theorem states that

$$P(k_y, k_z) = A(0, k_y, k_z) \tag{5}$$

Suppose, for example, that we are given three orthogonal image planes, $p_x(y, z)$, $p_y(x, z)$, and $p_z(x, y)$, whose in-plane pixels are all squares with a side of 1 unit. The three planes have thicknesses of N_1, N_2, and N_3 units, in the x, y, and z directions respectively. The intersection of these three planes is an $N_1 \times N_2 \times N_3$ unit right hexahedron. Without loss of generality, suppose that the hexahedron has corners at the points $(-N_1/2, -N_2/2, -N_3/2)$ and $(N_1/2, N_2/2, N_3/2)$. Let $P_x(k_y, k_z)$ be the discrete Fourier transform of the $N_2 \times N_3$ rectangle in $p_x(y, z)$ centered at the origin, and let $P_y(k_x, k_z)$ and $P_z(k_x, k_y)$ be similarly defined. Then a volumetric image $a(x, y, z)$ containing $N_1 \times N_2 \times N_3$ unit-cube voxels can be created by inverse Fourier transforming the following function:

$$\begin{aligned} A(0, k_y, k_z) &= P_x(k_y, k_z) && (6)\\ A(k_x, 0, k_z) &= P_y(k_x, k_z) && (7)\\ A(k_x, k_y, 0) &= P_z(k_x, k_z) && (8)\\ & && (9) \end{aligned}$$

In theory, the orthogonal plane measurements $P_x(0, k_z)$ and $P_y(0, k_z)$ are both measurements of $A(0, 0, k_z)$, so they should be equal. In practice, these values will be slightly different, due to thermal noise. It is therefore possible to reduce image noise slightly by averaging:

$$A(0, 0, k_z) = \frac{P_x(0, k_z) + P_y(0, k_z)}{2} \tag{10}$$

Notice that the only specified elements of $A(k_x, k_y, k_z)$ are those for which either $k_x = 0$, $k_y = 0$, or $k_z = 0$. Simultaneous variation in all three principal directions is not recovered, resulting in some orientation-dependent smoothing of the final image. In theory, the additional samples of $A(k_x, k_y, k_z)$ could be obtained by collecting additional image planes with appropriate spatial orientations, but in practice the extra cost makes this approach infeasible. The best solution to this problem, therefore, is to limit the number of missing samples by limiting the size of the Fourier transforms, either by limiting image plane thickness to a maximum of two or three units, or by smoothing and downsampling the projection images before 3D reconstruction.

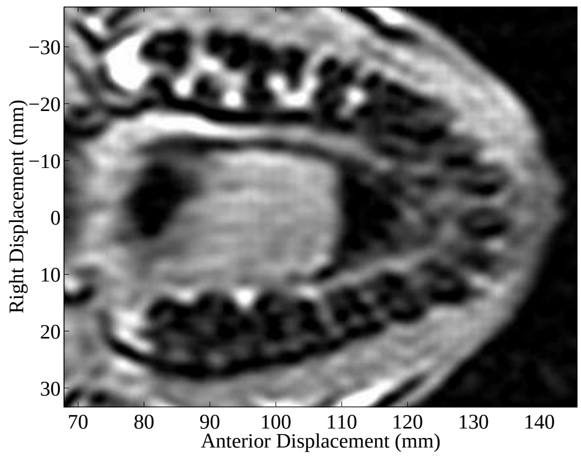

Figure 5: Axial image of the mouth, resampled to a uniform in-plane pixel size of $5/3 \times 5/3$mm. Slice thickness is 5mm.

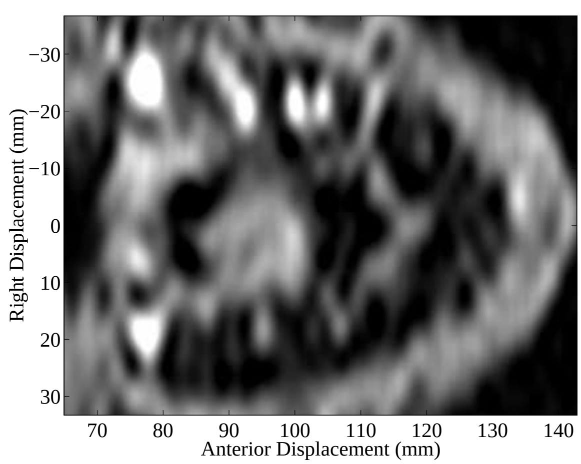

Figure 6: Axial slice through a stack of sagittal images. Voxel size is 5mm in the right-left direction, and 5/3mm in the anterior-posterior direction.

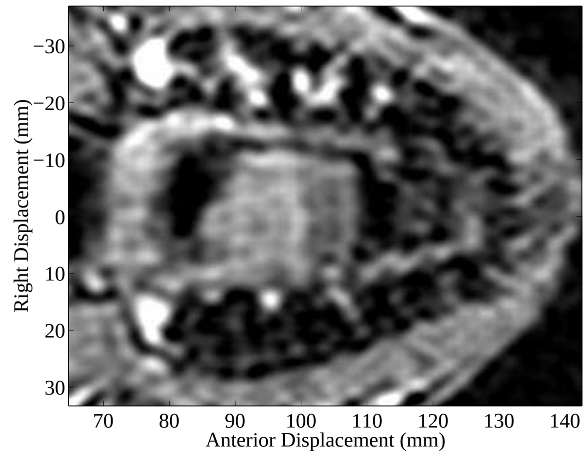

Figure 7: Axial slice through the reconstructed uniformly sampled image a(x,y,z). Voxels in the reconstructed image are cubes measuring 5/3mm on each side.

3 Experimental Results

The sagittal, coronal, and axial image stacks introduced in figure 4 were resampled so that every voxel measures $5/3 \times 5/3 \times 5$mm. Figure 5 shows an axial image from this stack; pixels in the image are squares measuring 5/3mm on each side (note that the image has been interpolated again for display in this figure, so that the effect of resampling may not be obvious). Figure 6 shows an axial slice through the sagittal image stack. Pixels in this image are still 5/3mm in the anterior-posterior direction, but they are 5mm in the right-left direction, resulting in the distortion visible in the image.

The resampled sagittal, coronal, and axial image stacks aligned in three dimensions, and combined using the algorithm proposed in equations 6 through 10. Each $3 \times 3 \times 3$-voxel cube of $a(x, y, z)$ was constructed by combining information from a 3×3 sagittal transform $P_x(k_y, k_z)$, a 3×3 coronal transform $P_y(k_x, k_z)$, and a 3×3 axial transform $P_z(k_x, k_y)$. In this way, nineteen of the twenty-seven values in the function $A(k_x, k_y, k_z)$ were specified. The missing eight values were simply zeroed out:

$$A(k_x, k_y, k_z) = 0 \text{ if } k_x \neq 0,\ k_y \neq 0,\ k_z \neq 0 \quad (11)$$

The function $A(k_x, k_y, k_z)$ was inverse transformed to obtain a uniformly sampled image $a(x, y, z)$, with cubic voxels measuring 5/3mm on each side.

Figure 7 shows an axial slice through $a(x, y, z)$. Comparison with figure 5 shows that the resolution of the reconstructed image in the axial plane

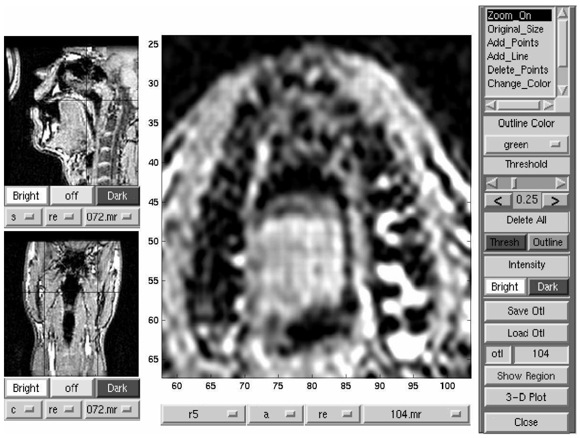

Figure 8: mrcat editing tool, showing sagittal, coronal, and axial slices from the reconstructed image stack a(x,y,z).

is exactly the same as the resolution of the axial image from which it was created. Notice, however, that there are some small artifacts in figure 7 which do not exist in figure 5, including for example the gray scalloped pattern on some of the edges of the tongue. These artifacts may be orientation-dependent smoothing artifacts, caused by the approximation defined in equation 11.

Figure 8 shows sagittal, coronal, and axial slices through $a(x, y, z)$, displayed in the mrcat viewing window. The main window is zoomed in to show a close up of the mouth area, and the locator images are also slightly zoomed. The images shown in figure 8 have an effective slice thickness of only 5/3mm, but there is no thermal noise visible in any of the images. In fact, with the exception of some possible orientation-dependent smoothing artifacts, the resolution of figure 8 looks as good as the resolution of the images with a 5mm slice thickness shown in figure 4.

4 Conclusions

This article has demonstrated an algorithm which creates a uniformly-sampled volumetric image by combining information from sagittal, coronal, and axial MR image stacks. In the example shown, a volumetric image $a(x, y, z)$ whose voxels are cubes measuring 5/3mm per side is created by combining image stacks with an original 5mm slice thickness. The level of thermal noise in the image $a(x, y, z)$ is as low as or lower than the level of thermal noise in any of the original image stacks. Unfortunately, information in the original image stacks is not completely sufficient to specify $a(x, y, z)$, and the image $a(x, y, z)$ may suffer minor, short-range orientation-dependent smoothing artifacts as a result. The importance of these artifacts depends on the application.

References

[1] Mark Hasegawa-Johnson, Jul Setsu Cha, and Katherine Haker. CTMRedit: a matlab-based tool for segmenting and interpolating MRI and CT images in three orthogonal planes. In *21st Annual Int. Conf. of the IEEE/EMBS Society*, Atlanta, GA, Oct. 1999.

[2] Mark Hasegawa-Johnson, Jul Setsu Cha, Shamala Pizza, and Katherine Haker. MR-CAT: A case study in human-computer interface design. In *Int. Conf. on Public Part. and Inform. Tech.*, Lisbon, Oct. 1999.

[3] Shrikanth S. Narayanan and Abeer A. Alwan. A nonlinear dynamical systems analysis of fricative consonants. *J. Acoust. Soc. Am.*, 97(4):2511–2524, 1995.

[4] John M. Ollinger and Jeffrey A. Fessler. Positron-emission tomography. *Signal Processing Magazine*, 14(1):43–55, January 1997.

[5] G. A. Wright. Magnetic resonance imaging. *Signal Processing Magazine*, 14(1):56–66, Jan. 1997.