Fusion of Short-Axis and Long-Axis Cardiac MR Images

A. Ardeshir Goshtasby, Ph.D.
Computer Science and Engineering
303 Russ Engineering Ctr
Dayton, OH 45435
ardeshir@cs.wright.edu

David A. Turner, M.D.
Diagnostic Radiology and Nuclear Medicine
Rush-Presbyterian-St. Luke's Med Ctr
Chicago, IL 60612
dturner@rad.rpslmc.edu

Abstract

A method is introduced for fusing the short-axis and long-axis cardiac MR images into an isotropic volume image. A volume image obtained by this method contains the left ventricular (LV) cavity in one piece, facilitating measurement of its shape and volume. The main goal in this image fusion is to reconstruct the LV cavity in volume form and in high resolution. The accuracy of the method is measured using a synthetic image. Examples of image fusion using real images are also presented.

1. Introduction

The accurate measurement of cardiac function requires that the volume and shape of the left ventricular (LV) cavity be accurately determined at different phases of the cardiac cycle [2] [7] [13] [18] [20] [22]. In the past, long-axis [6] [22] and short-axis [2] [7] [18] [20] [21] images have been used separately to measure the LV volume. Short-axis images provide high-resolution information about the shape of the LV at several cross-sections normal to the LV major axis while providing poor resolution normal to the axis. On the other hand, long-axis images provide high-resolution information at a few cross-sections parallel to the LV major axis while providing poor resolution normal to the axis. The objective in this work is to combine the long-axis and the short-axis images obtained under the same MR coordinate system and produce an image that provides high resolution information both along and normal to the LV major axis.

Reconstructing the LV cavity at a higher resolution enables the LV volume and shape to be measured more accurately. In the past, the ellipsoid has been used as a model to represent the LV cavity and estimate LV regions that are missing from given image slices. By fusing the long-axis and short axis images, we would like to reconstruct the entire LV cavity in high-resolution and measure its shape and volume without using a model. The shape of the LV cavity deviates considerably from an ellipsoid at some phases of the cardiac cycle. We will determine the shape of the LV cavity at a cardiac phase by segmenting the volumetric image obtained from the fusion of the long-axis and short-axis images at that phase.

To reconstruct the entire LV cavity at a particular cardiac phase, we map the long-axis and short-axis images obtained in that phase back into the “MR viewing window” (an orthogonal volume encompassing the imaged volume). This mapping determines some entries of the 3-D image, which represents the MR window. Other entries of the 3-D image are estimated by approximation. Once an isotropic volume image of the heart is obtained, the entire LV cavity will be extracted by segmentation.

In the following, first the existing methods that measure the LV volume using either the long-axis or the short-axis MR images are reviewed. Then, a method is introduced that measures the LV volume using both the long-axis and short-axis MR images. Finally, experimental results are given and evaluated using synthetic and real images.

2. Background

The left ventricular ejection fraction, a measure of cardiac function, is determined by measuring the LV volume at the end-systolic and end-diastolic phases of the cardiac cycle. Past methods have used either the short-axis or the long-axis images to determine the LV volume. Methods that are based on the long-axis images use a single image of the cross-section of the LV cavity that contains the LV major axis [22]. Assuming that the LV cavity is an ellipsoid, its cross-section with a plane that contains its major axis is an ellipse.
By fitting an ellipse to an obtained LV boundary, the parameters of the ellipsoid are determined. LV volume obtained by the estimated ellipsoid is then adjusted using experimental results [5]. Attempts to automate this method have been made by delineating LV boundaries using a computer program. Lilly et al. [15] developed a model-based approach to LV boundary extraction, while Duncan et al. [6] developed a method that was based on an energy-minimizing criterion.

Methods that are based on the short-axis images try to reconstruct the LV cavity first and then measure its volume [2] [7]. The challenge here is to acquire images that contain the entire LV cavity at the end-systolic and end-diastolic phases of the cardiac cycle, and also to accurately delineate the LV boundary. To determine the short-axis LV boundaries, Wang et al. [24] and Fleagle et al. [9] used interactive methods while Cohen [4] used an automatic method based on an energy minimizing process. Suh et al. [23] and Faber et al. [8] developed model-based methods for extraction of LV cross-sections.

An attempt has also been made to combine cardiac MR images obtained at transverse, coronal, and sagittal cross-sections. Kuwahara and Elho [17] first manually traced ventricular boundaries in two transverse, two coronal, and two sagittal cross-sections of the heart, and then combined the images to reconstruct the LV cavity in 3-D. In contrast to the method of Kuwahara and Elho, which combines LV boundary contours obtained at three orthogonal cross-sections of the heart, we combine the original short-axis and long-axis images (not their boundary contours) to construct a volumetric image of the LV cavity and the surrounding tissues. In addition, our method is automatic and does not require any manual tracing. The proposed method maps the long-axis and short-axis image slices back into 3-D. For each known pixel intensity in the image slices, it determines the corresponding voxel intensity in the 3-D image. It then estimates 3-D image intensities that are not obtained by this mapping through approximation. The details of the method are discussed below.

### 3. Method

Figure 1 shows a sequence of eight short-axis MR images of the heart, while Figure 2 shows a sequence of four long-axis images. These images represent cinematic (cine) MR images acquired with a 1.5-Tesla device (Signa; GE Medical Systems, Milwaukee). Imaging sections were 1 cm thick and contiguous; the matrix size was 256 × 256, and the field of view was 31 cm × 31 cm. The cine MR images were acquired with software provided by the manufacturer [10]. Blood that flows into the imaged volume during the acquisition of images with this technique emits signals of higher intensity than tissue that remains in the volume during the entire acquisition (e.g., ventricular myocardium), a phenomenon known as “flow enhancement.” Thus the blood pools appear brighter than other tissues in an acquired image. Figure 3a depicts the locations of the 2-D image slices of Figures 1 and 2 in the 3-D MR window. Figure 3b shows the 2-D images as they appear after being mapped into the window. This mapping partially determines the intensity of voxels in the 3-D image. It is possible that a pixel in a short-axis slice and a pixel in a long-axis slice map to the same voxel in 3-D. In such a situation, we let the intensity at the voxel equal the average of the two pixel intensities.

The equation describing the location of each MR slice in 3-D can be determined from the information provided in the header of the slice. By reading the 3-D coordinates of the upper-left, upper-right, and lower-left corners of each image slice from its header, the equation of the plane in 3-D from which the image was scanned can be determined. Assuming the coordinates of points in an image slice are represented by \((x, y)\) and the coordinates of the same points in 3-D are represented by \((X, Y, Z)\), the relation between the coordinates of corresponding points in 2-D and 3-D can then be written as:

\[
X = a_1 x + a_2 y + a_3 \tag{1}
\]

\[
Y = a_4 x + a_5 y + a_6 \tag{2}
\]

\[
Z = a_7 x + a_8 y + a_9 \tag{3}
\]

A 2-D image is obtained by translating, rotating, and scaling a cross-section of the 3-D window. This transformation is linear and can be represented by relations (1)-(3). There are nine unknown parameters in this transformation, \((a_1-a_9)\), which can be determined by substituting the coordinates of the upper-left, upper-right, and lower-left corners of an image slice in 2-D and their correspondences in 3-D into equations (1)-(3). The obtained system of equations can then be solved. Note that each image corner has three coordinates. When the coordinates of three corners are substituted into equations (1)-(3), nine equations will be obtained from which the nine unknowns are determined.

Once the unknown parameters of equations (1)-(3) are determined, and the coordinates \((x, y)\) of a pixel in an image slice are substituted into equations (1)-(3), the corresponding voxel position in the 3-D image can be computed. Relations (1)-(3) thus establish correspondence between pixels in a sequence of image slices and the corresponding voxels in the 3-D image. By
mapping the image slices to 3-D in this manner, we obtain the image shown in Figure 3b. The planes in the MR window corresponding to the long-axis and short-axis slices are shown in Figure 3a. Only a small portion of image 3b, as shown in Figure 4b, contains the LV cavity. We concentrate on the subwindow, which contains the LV cavity and discard the rest. The stripes in Figure 4b occur because the MR window was quantized into a 3-D array whose cross-section was slightly larger than its corresponding 2-D image slice. For voxels falling on the narrower stripes, no corresponding image pixels existed in the provided image slices. Figure 4a more clearly shows voxels where image intensities did not exist in the given image slices. We approximate the intensities of these voxels as well as the intensities of voxels that lie in between the image slices.

Suppose \( P = (X, Y, Z) \) is an image voxel whose intensity is unknown. In the following, we estimate the intensity at \( P \) using intensities of \( N \) voxels closest to it whose values are known. Let's suppose \( \{P_i : i = 1, \ldots, N\} \) are the coordinates and \( \{I_i : 1, \ldots, N\} \) are the intensities of the \( N \) voxels. We approximate the intensity \( I \) at voxel \( P \) using a weighted sum of intensities at the \( N \) voxels:

\[
I(P) = \frac{\sum_{i=1}^{N} I_i D_i(P)}{\sum_{i=1}^{N} D_i(P)}
\]

(4)

where

\[
D_i(P) = [(X - X_i)^2 + (Y - Y_i)^2 + (Z - Z_i)^2]^{-1}
\]

(5)

\( D_i(P) \) represents the inverse squared distance between voxels \( P \) and \( P_i \). The farther voxel \( P_i \) is from voxel \( P \), the smaller the contribution of its intensity to the estimated intensity at \( P \). And conversely, the closer voxel \( P_i \) is to \( P \), the larger its influence on the estimated intensity at \( P \). The square-root of formula (6) has been widely used as the basis function of multiquadrics in approximation theory [3] [12] [14] [19]. The basis functions in multiquadrics, however, decrease gradually from a center point and are not suitable for our purposes. In our formulations, we use the square distances, which decrease more rapidly from a center point and are computationally faster also. We introduce a smoothness parameter \( r^2 \) into equation (5) as is done in the basis functions of multiquadrics, to obtain:

\[
D_i(P) = [(X - X_i)^2 + (Y - Y_i)^2 + (Z - Z_i)^2 + r^2]^{-1}
\]

(6)

As parameter \( r^2 \) is increased the contribution from farther voxels on an estimated intensity increases, thus producing a smoother approximation. However, as \( r^2 \) is decreased, the contributions increase from the immediate neighbors of the voxel whose intensity is being estimated. Thus a more detailed approximation is produced.

Although \( N \) can represent the total number of voxels whose intensities are known in a 3-D image, in practice however, the estimated intensity at voxel \( P \) depends mainly on the intensities of a small number of voxels neighboring \( P \). As one of these voxels moves away from \( P \), the effect of the image voxels on the estimated intensity vanishes rapidly. The neighborhood size to be used in this approximation depends on the distance between the given image slices and also on parameter \( r^2 \). As the distance between image slices increases, larger neighborhoods should be used to allow a sufficiently large number of voxels to participate in the approximation. Also, as \( r^2 \) increases, the basis function used in the approximation becomes wider, covering a larger portion of the image and requiring a larger neighborhood for the approximation. For the integer distance of \( d \) pixels, we have found that the windows of size \( w \times w \times w \), where \( w = (2dr^2 - 1) \), must be used to achieve an approximation error less than half a pixel.

For instance, when \( r^2 = 1 \) and \( d = 8 \), we need windows of size \( 15 \times 15 \times 15 \), when \( r^2 = 2 \) and \( d = 8 \), we need windows of size \( 31 \times 31 \times 31 \), and so on. Voxels falling outside of these windows have no effect on an estimated intensity with digital accuracy.

Applying the approximation of formulas (4) and (6) to the data of Figure 4b, we obtain the image shown in Figure 4c (only cross-sections of the volumetric image with its bounding planes are depicted). In the computations, we used \( r^2 = 1 \). Oblique cross-sections of this image are shown in Figure 5 (here also only the cross-sections of the volumetric image with its bounding planes are shown). In the following, we determine the accuracy of the approximation by reconstructing a synthetic volume image whose geometry is known, from its cross-sections.

4. Evaluation

In this section to determine the accuracy of a reconstructed volume image from given short-axis and long-axis image slices, we use an image that contains an object whose geometry is known. A \( 128 \times 128 \times 128 \) synthetic image containing an object composed of a cylinder, a cube, a cone, and a sphere was generated. The intensity of voxels belonging to the object was 200 while that of the background was 100. Figure 6a depicts the obtained object after thresholding the synthetic image at 150. The synthetic image was partitioned into \( 128 \times 128 \times 8 \) subimages, and each subimage was reduced to a \( 128 \times 128 \) image by averaging the eight voxels along the Z-axis into a pixel value in
the 2-D image. Next, the 3-D image was partitioned into 128 \times 8 \times 128 subimages, and each subimage was reduced to a 128 \times 128 image in the same fashion.

Assuming that the obtained 2-D images represented MR slices obtained at two orthogonal scan directions, we then fused them together using the method described in the preceding section. The reconstructed image is shown in Figure 6b. Parameter \( r^2 \) was equal to 0.5 in this approximation. On increasing \( r^2 \) to 1.0 and 2.0, we obtained the results shown in Figures 6c and 6d, respectively. As parameter \( r^2 \) is increased, the approximation neighborhood becomes larger, producing a smoother approximation. These figures show the surfaces obtained after thresholding the reconstructed images at 150.

The root-mean-squared difference between intensities of Figure 6a and intensities of Figures 6b-d were 7.45, 6.50, and 8.64, respectively. As parameter \( r^2 \) is increased, intensities in large gaps are determined more accurately. As \( r^2 \) is decreased, intensities near known voxels become more accurate. Among these three cases, average error was the smallest when \( r^2 = 1.0 \).

As interslice distance decreases, a denser distribution of voxels whose intensities are known becomes available. In such a situation, a smaller \( r^2 \) should be used so the details in the image are not smoothed out. As interslice distance increases, a larger \( r^2 \) should be used to allow a sufficiently large number of voxels to participate in estimating the intensities of large gaps.

Figure 7 shows the blood pools of Figure 4c obtained by the marching cubes algorithm \([16]\). The threshold value used in the algorithm was determined by interactively varying the intensity in one of the image slices until the best LV boundary was obtained by a radiologist’s judgement. Then, the same threshold value was used in the marching cubes algorithm to determine the blood pools. More elaborate segmentation techniques \([1] [4] [8] [11]\) may also be used to extract the blood pools once the volume image is constructed.

Parameter \( r^2 \) was equal to 0.5 in Figures 7a-b. On increasing \( r^2 \) to 1.0 and 2.0, we obtained the images shown in Figures 7c-d and 7e-f, respectively. As parameter \( r^2 \) is increased, a smoother approximation is obtained. A smoother approximation reduces image noise. At the same time, however, image details are also smoothed out. As can be observed in Figure 7, as \( r^2 \) is increased, details on extracted ventricular surfaces reduce. A second example is shown in Figure 8 using the same parameters as in Figure 7. We see that in Figures 7 and 8 the entire LV cavity and the left atrium have been extracted. Most parts of the right ventricular cavity has also been extracted. The short-axis and long-axis images used in these experiments did not contain the right atrium and, therefore, the right atrium was not obtained in the reconstructed images. The important point to note in these images is that the fusion process produces images that contain the entire LV cavity in high resolution. The process may be repeated on images from different phases of the cardiac cycle to extract the LV cavity in its entirety, thus enabling measurement of dynamic shape and volume of the LV cavity.

5. Summary and Conclusions

Measurement of volume-based indices of cardiac function requires measurement of the LV volume and shape at different phases of the cardiac cycle. Since the heart moves during the cardiac cycle, no single 2-D image location contains the same part of the heart during the entire cycle. Therefore, in order to follow changes in cardiac volume and shape through the entire cycle, the image planes must be interrogated to include all parts of the heart during its entire cycle. When conventional 2-D imaging techniques are used, the relatively poor spatial resolution along the dimension normal to the image planes distorts the apparent boundaries of the cardiac chambers along that dimension. This is particularly problematic in short-axis views, since the very thin cardiac valves cannot be adequately represented without excellent spatial resolution normal to the cardiac valve plane. However, spatial resolution along the normal dimension can be recovered by acquiring a second set of images parallel to that dimension (i.e., normal to the first set of images) and fusing the two sets of 2-D images in a single 3-D representation, as we have done.

The method introduced for fusion of the long-axis and short-axis MR images involves mapping the image slices back into the MR viewing window and estimating image entries that are not obtained by this mapping. The approximation process used in this work is based on a weighted averaging scheme with weights inversely proportional to the distances of voxels with known intensities to the voxel whose intensity is being estimated. This approximation has a smoothness parameter that can be varied to fuse images with different interslice distances.

Computing an entry of the estimated volume image from given image slices requires in the order of \( N \) multiplications, where \( N \) is the number of voxels whose intensities are known in a small neighborhood of the entry. Neighborhood size \( w \times w \times w \) is obtained from the interslice distance \( d \) and the smoothness parameter \( r^2 \) of the approximation: \( w = 2dr^2 \ - 1 \). Assuming the
reconstructed image is of size $n \times n \times n$, the algorithm requires in the order of $Nn^3$ multiplications to construct an isotropic volume image from given long-axis and short-axis image slices. Our implementation of the algorithm on a Sun Sparstation 10 required about one hour and forty-five minutes to obtain the volume image of Figure 4c (140 x 100 x 178 voxels) from image slices of Figures 1 and 2 (256 x 256 pixels).

References


Figure 1. A sequence of short-axis MR images of the heart obtained at the end-diastolic phase.

Figure 2. A sequence of long-axis MR images obtained at the end-diastolic phase.
Figure 3. The MR viewing window. (a) Planes in the MR window corresponding to the short-axis and long-axis images of Figures 1 and 2. (b) The short-axis and long-axis images after being mapped back to the 3-D MR window.

Figure 4. (a), (b) Portions of Figures 3a and 3b, respectively, containing the heart. (c) Volumetric image obtained from image of (b) using the approximation formulas of (4) and (6).
Figure 5. (a), (b) Two oblique cross-sections of the image of Figure 4c.

Figure 6. (a) A synthetic image containing an object composed of a sphere, a cube, a cylinder, and a cone. This image was reduced to a sequence of sixteen image slices parallel to the $xy$-plane and sixteen slices parallel to the $xz$-plane. The two image sequences were then fused together by the method described in the paper with $r^2 = 0.5$, 1.0, and 2.0 to obtain images (b), (c), and (d), respectively.
Figure 7. The blood pools when viewed from (a) the cardiac apex, and (b) from the left posterior towards the right anterior of the heart. These images were obtained when parameter $r^2$ in formula (6) was equal to 0.5. (c)–(d) and (e)–(f) are the same as (a)–(b), except that parameter $r^2$ was increased to 1.0 and 2.0, respectively.
Figure 8. Another example showing the reconstructed blood pools after fusing the long-axis and short-axis MR images from a second patient. Parameters of the image fusion are the same as those used in Figure 7.