

Automatic Volumetric Analysis of the Left Ventricle in 3D Apical Echocardiographs

by

Andrew James Wald

Department of Biomedical Engineering
Duke University

Date:

Approved:

Olaf T von Ramm, Supervisor

Joseph Kisslo

Fan Yuan

Sina Farsiu

Thesis submitted in partial fulfillment of
the requirements for the degree of
Master of Science in the Department of
Biomedical Engineering in the Graduate School
of Duke University

2015

ABSTRACT

Automatic Volumetric Analysis of the Left Ventricle in 3D Apical Echocardiographs

by

Andrew James Wald

Department of Biomedical Engineering
Duke University

Date:

Approved:

Olaf T von Ramm, Supervisor

Joseph Kisslo

Fan Yuan

Sina Farsiu

An abstract of a thesis submitted in partial
fulfillment of the requirements for the degree
of Master of Science in the Department of
Biomedical Engineering in the Graduate School of
Duke University

2015

Copyright by
Andrew James Wald
2015

Abstract

Apically-acquired 3D echocardiographs (echoes) are becoming a standard data component in the clinical evaluation of left ventricular (LV) function. Ejection fraction (EF) is one of the key quantitative biomarkers derived from echoes and used by echocardiographers to study a patient's heart function. In present clinical practice, EF is either grossly estimated by experienced observers, approximated using orthogonal 2D slices and Simpson's method, determined by manual segmentation of the LV lumen, or measured using semi-automatic proprietary software such as Philips QLab-3DQ. Each of these methods requires particular skill by the operator, and may be time-intensive, subject to variability, or both.

To address this, I have developed a novel, fully automatic method to LV segmentation in 3D echoes that offers EF calculation on clinical datasets at the push of a button. The solution is built on a pipeline that utilizes a number of image processing and feature detection methods specifically adopted to the 3D ultrasound modality. It is designed to be reasonably robust at handling dropout and missing features typical in clinical echocardiography. It is hypothesized that this method can displace the need for sonographer input, yet provide results statistically indistinguishable from those of experienced sonographers using QLab-3DQ, the current gold standard that is employed at Duke University Hospital.

A pre-clinical validation set, which was also used for iterative algorithm development, consisted of 70 cases previously seen at Duke. Of these, manual segmentations of 7 clinical cases were compared to the algorithm. The final algorithm predicts EF within ± 0.02 ratio units for 5 of them, and ± 0.09 units for the remaining 2 cases, within common clinical tolerance. Another 13 of the cases, often used for sonographer training and rated as having good image quality, were analyzed using QLab-3DQ, in which 11 cases showed concordance (± 0.10) with the algorithm. The remaining 50 cases retrospectively recruited at Duke and representative of everyday image quality showed 62% concordance (± 0.10) of QLab-3DQ with the algorithm. The fraction of concordant cases is highly dependent on image quality, and concordance improves greatly upon disqualification of poor quality images. Visual comparison of the QLab-3DQ segmentation to my algorithm overlaid on top of the original echoes also suggests that my method may be preferable or of high utility even in cases of EF discordance. This paper describes the algorithm and offers justifications for the adopted methods. The paper also discusses the design of a retrospective clinical trial now underway at Duke with 60 additional unseen cases intended only for independent validation.

Index Terms— echocardiography, ultrasound, ejection fraction, automatic, segmentation, image processing

Dedication

This thesis is dedicated to Claude Shannon, the father of information theory and electrical engineering, and a true engineering innovator whose legacy is carried in all mankind's worthiest endeavors. You, sir, are my inspiration.

Contents

Abstract	iv
List of Tables	ix
List of Figures	x
Acknowledgements	xiii
1. Introduction	1
1.1 Background	1
1.2 Overview	2
2. Method & Algorithm Design	4
2.1 Segmentation Orientation	5
2.2 Depth Filtering.....	6
2.3 Short-Axis Filtering.....	8
2.4 Iterative Centroid Estimation	9
2.5 Feature Detection.....	11
2.6 Ventricle Wall Census.....	15
2.7 Segmentation Optimization	18
2.8 Slice Area Calculation.....	22
2.9 LV-LA Depth Delineation: Method 1	22
2.10 LV-LA Depth Delineation: Method 2	24
2.11 LV Volume & EF Calculation.....	30
3. Results & Pre-Clinical Validation	35

3.1 Comparison to Manual Segmentation.....	35
3.2 Comparison to QLab-3DQ	37
4. Clinical Trial Design	41
5. Conclusion	43
References	44

List of Tables

Table 1. Comparison of ejection fractions for 7 training cases as determined by the algorithm to those obtained from manually segmented volumes.....36

Table 2: Comparison of ejection fraction as determined by algorithm to that determined by QLab-3QD as used by an experienced sonographer and/or Wald.....38

List of Figures

Figure 1: Overview of algorithm pipeline	4
Figure 2: Long-axis slices require segmentations of complex geometry (left) and decisions on whether to exclude or include papillary muscles in the segmented areas (mid-left). The segmentation becomes very difficult as one approaches the anterior wall, as many disjoint areas need to be produced (mid-right). Short-axis slices have much simpler geometry but require inclusion of chordae tendineae, which can falsely appear very bloated in some frames due to motion (right).	6
Figure 3: Issues frequently encountered in short-axis slices from 3D US apical datasets. The selected examples show arbitrariness of the ventricle wall (top left and middle). Poor resolve of significant sectors of the wall (top right, bottom left), confounding artifacts (bottom middle) and missing wall along with gating asynchrony (bottom right).	7
Figure 4: Two examples showing the effect on the original short-axis slices (left) after depth-filtering (middle) and 2D slice filtering (right). The 1D Blackman window used depth-wise preserves depth-dependent information that would otherwise be unused. The 2D slice filters remove speckle and increase contrast between fluid and tissue, necessary to detect the relevant tissue boundaries.	9
Figure 5: The long-axis estimation method locates points that are well within the ventricle (purple-starred centroids) even in these three difficult selections. Examples of three short-axis slices where the ventricle’s boundaries are poorly defined (left and middle) or arbitrary (right).....	11
Figure 6: Example radial search and feature detection using all 12 filters to identify potential boundaries. Note how all filter variants are necessary to detect all the features ultimately required to draw the endocardial boundary. Any one filter on its own does not detect all the correct features of interest.	14
Figure 7: After the radial search features are merged to form census radii for each angle, outliers are removed from the radii series as per my protocol, and then curve-fitting through the radii series is performed. The algorithm proceeds with the medians of these curve fits.	17
Figure 8: The median curve fit (white) is passed through the optimizer to produce another curve (purple). These selected examples show the optimizer adds small	

perturbations that improve the integrity of the segmentation. Finally, the radius of each angle is extracted from the optimized fit, and a 1D smoothening filter is used in the depth dimension on each angle to produce the final curve fit (peach).....	21
Figure 9: For one selected frame, the sector radii are shown along with the average radius (magenta) at every depth through the sample (top). The area of the ventricle at every depth, as segmented by the curve before the optimizer is used (red), post-optimization (green) and post-filtering (blue), is shown (middle). The final search region for the boundary between the atrium and ventricle is delimited to the same region for all frames, and the maximum of the moving average within this region of the deviation of the unfiltered ventricle center estimate from the long axis estimate is used to choose the separation depth (bottom).	24
Figure 10. Histologically stained section (right) shows the fibrous skeletal feature of the mitral annulus, with correlation to the same feature in the 4-chamber image (left). Credit: Anderson, R. & Becker, A.: The Cardiac Skeleton. Gower Medical Publishing, Ltd. 1980.	25
Figure 12. Long-Axis Filtering prior to feature detection for mitral annulus detection. Top: After a 1D Blackman impulse response is applied in the height dimension (face the page). Bottom: 2D filtering (same algorithm as Section 2.3) by contrast adjustment, Wiener filter, and order statistic filter.....	26
Figure 13. Feature detection method applied on the short axis is also used for the 4-chamber long-axis slices.....	27
Figure 14. Example of centroid (marked by star) located on a column-wise slice (4-chamber, even though only 2 chambers may be clearly visible).....	28
Figure 15. Points on the mitral annulus possibly identified after filtering the feature detection output on the long axis. The ventricle center is shown along with six points of interest.	29
Figure 16. Example of the LV-LA delineation being approximated by an estimated determination of the depth of the mitral annulus, for a particular frame, shown on top of the 4-chamber view.....	30
Figure 17: Volume time series for two cases. The output of the algorithm is compared to a manual segmentation.	31

Figure 18: The boundary of the segmented lumen can be viewed in context with the rest of the anatomy using a 3D volume rendering. This methodology was part of the verification process, and trivially implemented using the segmented volume. 32

Figure 19: (Top) Two orthogonal planar selections of a dataset's EDV with segmentation boundaries overlaid. (Bottom) Another two orthogonal planar selections of a different dataset also in EDV with segmentation boundaries shown. 33

Figure 20: (Top) Two orthogonal planar selections of a dataset's ESV with segmentation boundaries overlaid. (Bottom) Another two orthogonal planar selections of a different dataset also in ESV with segmentation boundaries shown. 34

Acknowledgements

I first wish to thank my engineering supervisor, Olaf von Ramm, who provided me with the opportunity and mentorship to pursue this project, while letting me do it my way. I will never forget Olaf's thought-provoking lectures.

I also wish to thank Joseph Kisslo, my clinical supervisor, without whose clinical insight this project would not have been motivated or possible. I will never forget the fun I had staring at clinical cases and being asked, "So, what's wrong with this one?" Joe has helped me understand the difficult realities of trying to be an innovator in this field.

Cooper Moore helped me in myriad ways and convinced me that this project could turn into a worthwhile thesis. He is an expert at precision pipetting-by-mouth, a skill which came in handy.

I appreciate Fan Yuan and Sina Farsiu for sitting on my committee and providing feedback without much forewarning.

Lastly, I wish to acknowledge the PGS-M grant awarded to me by NSERC, which gifted me the opportunity to pick Duke BME, a program which would have otherwise been out of my reach.

1. Introduction

1.1 Background

The left ventricle (LV) is the body's main pump for systemic circulation. Two of the key physiological biometrics used in the assessment of overall heart function by cardiologists are the LV stroke volume and ejection fraction (EF). To calculate these, volumetric analysis of the LV between end diastole (ED) and end systole (ES) are essential. Time series of 3D echocardiographs (echoes) through the cardiac cycle acquired in the standard 3- or 4-chamber apical view are increasingly being used to produce these volumes.

Echocardiographers may grossly estimate EF by visual inspection, or may use Simpson's or other methods to quickly produce an approximation of volumes. More accurate methods require segmentation of the LV lumen from the endocardial border. If done manually, the task is tedious, time-consuming, and error-prone. Semi-automatic methods in commercial systems, such as that used in the proprietary Philips Xcelera® QLab-3DQ (Andover, MA), have become standard over the last decade but are still prone to variability and require operator skill.

There have been a small number of studies aiming to automate the process over the last decade, such as [1-6], none of which have been pervasive clinically. These methods either require human input, and hence are dependent on operator judgment,

were developed using ideal images unrepresentative of the clinical reality, or were found ineffective on clinical cases observed at Duke. The latter was especially the case for diseased patients whom also tend to produce images featuring high signal drop-out.

This study was clinically motivated by the current non-existence of a completely automatic, deterministic, and robust method to corroborate and ideally replace existing methods for quantifying LV volumes and EF using apically-acquired 3D echoes.

1.2 Overview

This paper present a novel, fully automatic, push-button solution to the segmentation problem involved in LVEF acquisition using apical 3D echoes. The method may augment existing commercial systems and is intended for clinical application. This paper describes the algorithm and provides design justifications, highlighting the practical issues encountered with real clinical datasets.

The algorithm is an image processing pipeline based on a number of filtering and feature detection methods that have been specifically adopted to handle the variable quality, incompleteness and speckle inherent in the 3D echo modality. In brief, the algorithm attempts to accommodate for commonly missing backscatter contrast of endocardium by leveraging the 3D nature of the dataset and general assumptions of LV geometry. The short-axis slices are filtered to remove speckle, increase contrast, and emphasize tissue boundaries. The long axis is estimated using a converging centroid

estimator on each short axis. Feature detection is performed radially about the long axis using an ensemble of blurring and gradient magnitude filters. These points are used to fit estimates of the endocardial border about the long axis. An optimization method enhances this border, and the 3D nature of the dataset is again leveraged to inform a luminal shell. The LV is delineated from the left atrium by taking advantage of the predictably high contrast of the fibrous skeleton of the mitral annulus.

The algorithm was iteratively designed using feedback on a set of 70 training cases acquired at Duke, which doubled as a preclinical validation set. It was found that success may be highly dependent on image quality, and that disqualification of poor quality candidates lends to an optimal niche. Based on positive findings presented, my group decided to proceed with independent clinical validation on a set of 60 additional cases. This retrospective clinical trial has been designed to address the hypothesis that the method described in this paper can substitute for semi-automatic application of QLab-3DQ, the current gold standard at Duke University Hospital, in the determination of EF without yielding results of clinically significant difference. This paper will also discuss the design of this trial.

2. Method & Algorithm Design

An overview of the steps of my methodology is provided in Figure 1. My method segments the LV and outputs a binary, 3D matrix giving a Cartesian volume model for each captured frame independently. Assuming the operator has captured diastole and systole, a variety of analyses can then be carried out on the time series, notably the calculation of EF. It is assumed that the operator has access to the post-processed, scan-converted Cartesian volumes in DICOM format, which are the only inputs needed.

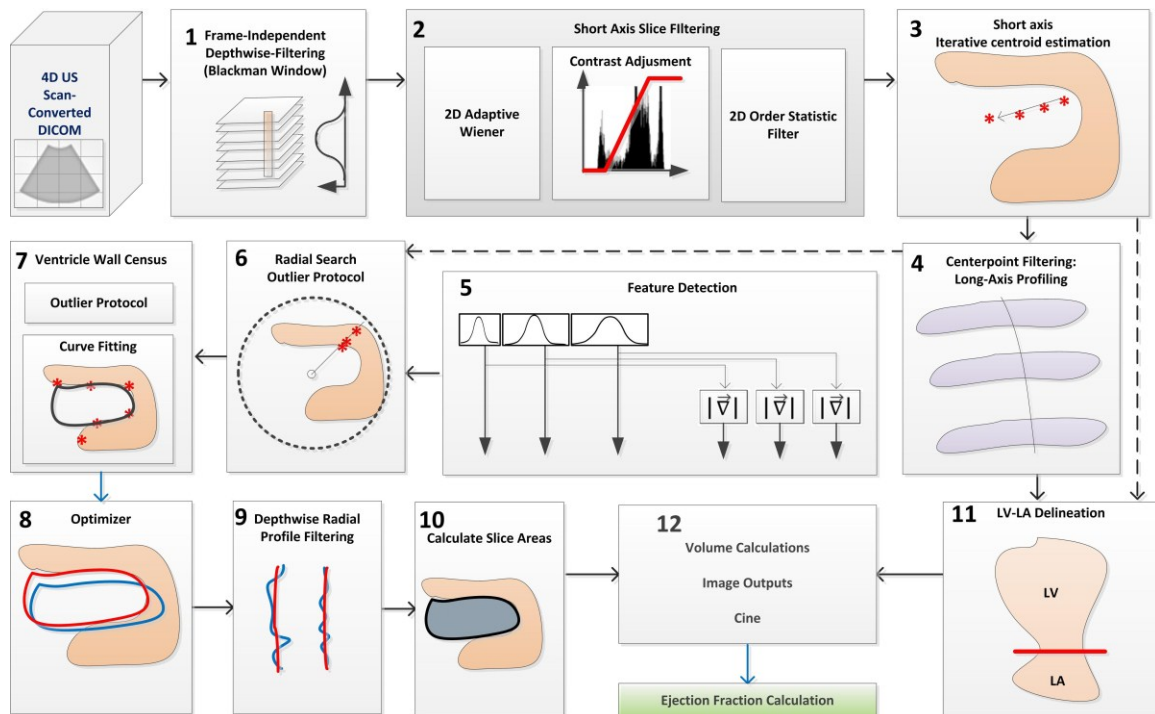


Figure 1: Overview of algorithm pipeline

2.1 Segmentation Orientation

Practical issues with 3D apical data learned from all 7 datasets were considered during design. My first design decision was whether to segment the 3D volume directly or to acquire a series of 2D areas and use those to produce a volume. After studying many short- and long-axis slices, it was concluded that 2D short-axis filtering was geometrically the simplest problem to solve. Consider Figure 2, which shows a few representative long-axis slices, including one near the posterior wall. The coarseness of the ventricle wall at the posterior would necessitate the fit of a relatively complex geometry for both the 2D long-axis and 3D views. While this coarseness is still present in 2D short-axis views, simpler shapes with fewer degrees of freedom would clearly work well for approximating area. Hence, I developed an algorithm to do slice-by-slice segmentations on the short-axis, with heavy cross-dependence of adjacent segmentations, as will be described. A radial profile of the ventricular wall about the long-axis is made for every short axis slice between the apex and base, yielding a calculation of the area of the LV as intersected by the slice plane. Integrating these areas through slice thickness over the range of apex to base gives the volume estimate.

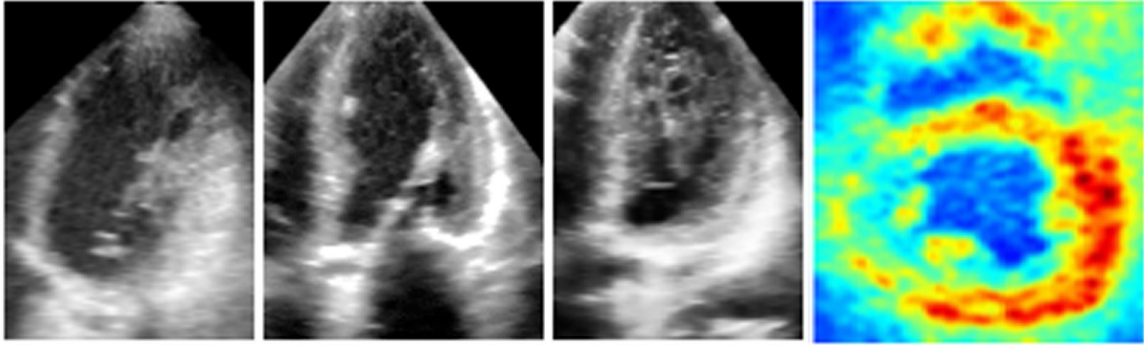


Figure 2: Long-axis slices require segmentations of complex geometry (left) and decisions on whether to exclude or include papillary muscles in the segmented areas (mid-left). The segmentation becomes very difficult as one approaches the anterior wall, as many disjoint areas need to be produced (mid-right). Short-axis slices have much simpler geometry but require inclusion of chordae tendineae, which can falsely appear very bloated in some frames due to motion (right).

2.2 Depth Filtering

Poor backscattering signal and drop-out can arise for a number of reasons, including malfunctioned elements on the transducer, poor penetration and excess attenuation at the frequencies used. A feature present to varying extent in most of my datasets was the invisibility or non-differentiability of the ventricle wall for a subset of the slices. Figure 3 shows a number of examples where luminal extent is uncertain.

In some continuous depth windows, some over 0.5 cm long, up to half of the circumference about the long-axis is mostly missing backscatter contrast. Smaller sectors of missing wall contrast, measurable in the degrees to tens of degrees, occur randomly throughout all the images at a broad range of depths. To prevent my algorithm from failing at the boundary detection stage on poorly imaged sectors, the data is one-

dimensionally filtered in the depth dimension. Every stack of voxels undergoes an independent convolution with a Blackman window. Other apodization functions used for windowing, such as Hanning and Gaussian windows, achieved similar effect. Where this does not completely fix the missing wall problem in some patients, later stages in the pipeline do.

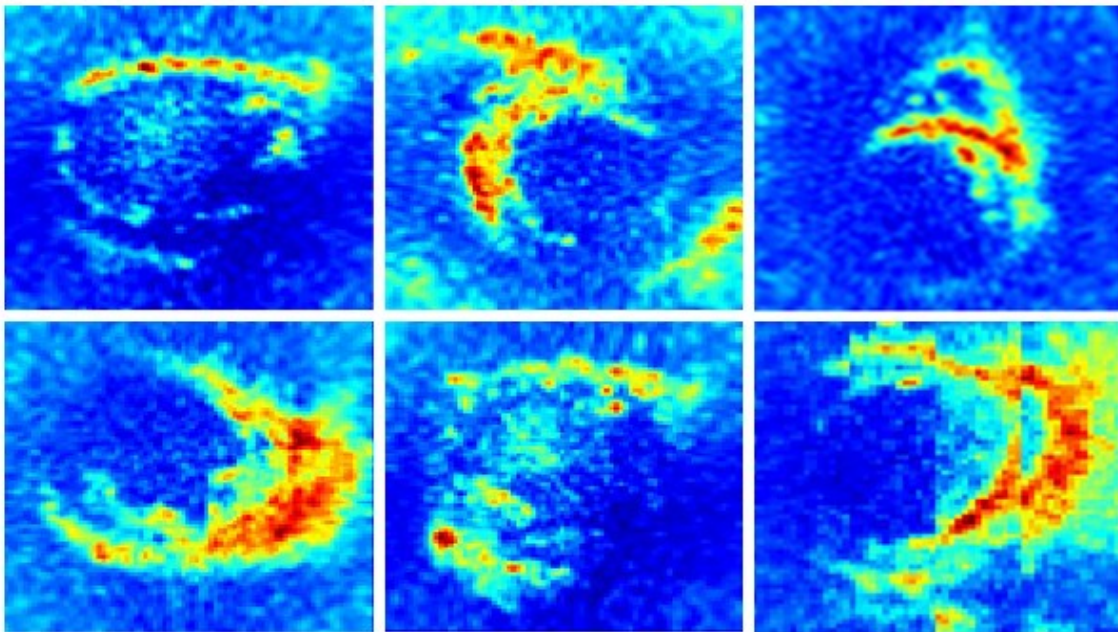


Figure 3: Issues frequently encountered in short-axis slices from 3D US apical datasets. The selected examples show arbitrariness of the ventricle wall (top left and middle). Poor resolve of significant sectors of the wall (top right, bottom left), confounding artifacts (bottom middle) and missing wall along with gating asynchrony (bottom right).

2.3 Short-Axis Filtering

Speckle is a random, constructive and destructive interference process that arises from the acoustic radiation reflections by sub-resolution scatterers throughout the medium. Different biomaterials have varying degrees of speckle. To some extent, the phenomenon ensures that material differences on a coarse scale are detected. On the other hand, speckle causes the appearance of more boundaries delineating the material types than those that actually exist. For successful detection of candidate wall points in the Feature Detection stage, it was found vital to suppress speckle and other noise processes on the short-axis slices. The 2D adaptive Wiener filter [7] was selected as a first sub-stage. To focus detection on the blood-endocardial boundary, the second sub-stage is a contrast adjustment that clips the lower 20% and upper 70% of dynamic range, mapping voxels beyond these thresholds to constant values. This considerably assists the success of edge detection in successive stages by removing bulk blood and bulk myocardium as distractors. The third stage is a 2D order-statistic filter using a moving window. Figure 4 shows the effect after depth-filtering and after the 3 filters described in this section. The parameters for each of these 3 stages were found to be flexible. All 3 of these filters cause resolution loss. However, given the scale of the LV, any such loss would likely have negligible effect on volume quantification.

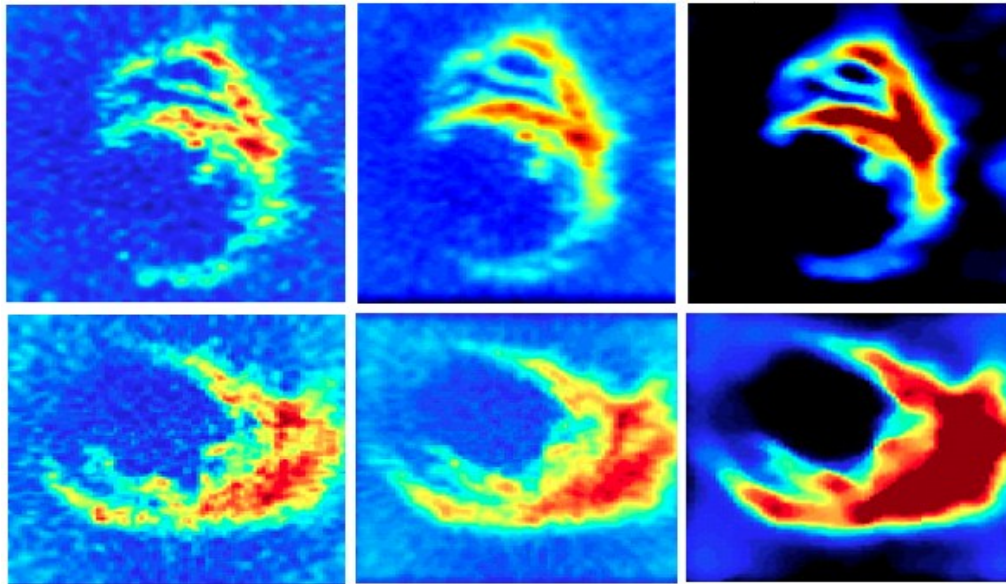


Figure 4: Two examples showing the effect on the original short-axis slices (left) after depth-filtering (middle) and 2D slice filtering (right). The 1D Blackman window used depth-wise preserves depth-dependent information that would otherwise be unused. The 2D slice filters remove speckle and increase contrast between fluid and tissue, necessary to detect the relevant tissue boundaries.

2.4 Iterative Centroid Estimation

As the latter stages of the algorithm require a long-axis estimate, an iterative estimation method was developed that converges on a centroid approximation for each short-axis slice. On the first iteration, the centroid is assumed to be in the middle of the frame. Even if this point is not inside the LV, it was found that all slices of all 7 patients ultimately do converge into the ventricle. The circumference is divided into at least 30 angles, each representing a sector. A radial line profile emanating from the centroid is drawn along each angle. For each angle, I search for the radius at which the largest

backscatter intensity occurs. In case of ties, I take the smallest radius. Once complete for all angles, I find the centroid of these points. The algorithm is then repeated about the new centroid. Again, a second centroid is found, and I repeat yet again. After some number of iterations, the centroid converges. The stopping criteria is when the centroid coordinates do not change between two successive iterations. While the vast majority of slices are stable and do converge, there are a significant number that are unstable and do not. Without an additional stopping criteria, this would cause the algorithm to run forever. Hence, the number of iterations is capped at 30. Typical, stable slices converged within 8 iterations, and over 90% within 4 iterations. On non-converging slices, I replace the centroid with the center of the frame as a best guess.

All frames for all depths are then squeezed into two 1D arrays for x and y of the centroid as found by the iterative method. For example, one would lay out frames 1 through 20 for the depth 1 slice, frames 1 through 20 for depth 2, frames 1 through 20 for depth 3, etc. Curve fitting is performed through these x and y series using robust moving-window regression (LOWESS) [8]. The smoothed centroid series connect to form a working approximation of the long axis. A window on the order of 30% of the squeezed series length is used. Inspection of the centroid estimation process revealed that we could successfully locate ourselves within the ventricle for even the worst-imaged patient. Examples of 3 short-axis slices where the LV's boundaries are poorly

defined, and yet the algorithm chooses correct estimations for points safely within the ventricle, are shown in Figure 5.

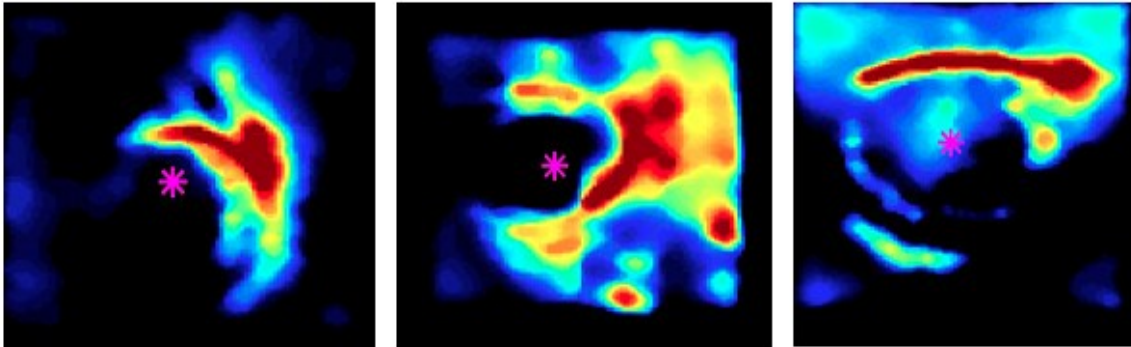


Figure 5: The long-axis estimation method locates points that are well within the ventricle (purple-starred centroids) even in these three difficult selections. Examples of three short-axis slices where the ventricle's boundaries are poorly defined (left and middle) or arbitrary (right).

2.5 Feature Detection

Looking radially from points known to be within the ventricle, closest-radius-max-intensity searching of neither voxel intensity nor magnitude gradient were adequate for identifying the blood-endocardial interface. Fine features, such as chordae in motion or unknown artifacts, would distract from the actual boundary despite the pre-filtering. Little-differentiated contrast of myocardium in some regions of the field aggravates the problem. The coarse ventricle wall often produces a margin of thickness where intensity contrast appears in-between that of blood and tissue. However, the critical observation was that different variations of pre-blurs with and without

subsequently taking the gradient magnitude offered partial selection of the wall boundary.

To leverage the combined information from these variants, the slices are convolved in parallel with a filter array. First, 6 different Gaussian blur point spread functions (PSFs) are convolved with every slice. Then, each blurred slice is passed through horizontal and vertical Sobel filters, whose outputs are added in quadrature to yield the magnitude gradient field of each blurred slice. The six Gaussian-blurred slices are also considered without the Sobel filters. Hence, each of the total of 12 different outputs has a particular combination of σ and whether or not the magnitude gradient filter was applied.

Varying the σ parameter of the Gaussian blur PSF has major impact on the types of edges detected. More blur (higher σ) means coarser-scale biomaterial changes are detected. For example, chordae may be washed out and would go unnoticed. However, tiny hints of ventricle wall also get washed out. Less blur (lower σ) increases the susceptibility of highlighting uninteresting features like chordae. Equal intervals, found by considering all 7 datasets, were chosen for σ with a starting value of zero (no blur).

The circumference of the short-axis slice about the long-axis is broken into a number of sectors. Considering as few as 40 different angles for the radial search were sufficient to converge on volume later. The radial search is repeated on each of the 12 slice versions out of the filter array. Along each sector angle's line profile, the closest

radius at which the maximum image intensity occurs along the line is selected. A few extra rules are placed in the protocol to improve the quality of points selected. During the radial search, the candidate points are concurrently analyzed in the pre-array slice. They are eliminated as distractors if the intensity of the same point in the pre-array slice is not in an “uphill” region, i.e. candidate for being on the boundary of interest. Points are also eliminated if their intensities in the pre-array slice are beyond the upper half of intensities throughout the slice, which will be occupied by bulk myocardium. This ensures no points are picked that are clearly inside the bulk muscle mass. The search diameter is limited to partially reduce the effect of irrelevant anatomy. Zero-intensities along the entire search profile of an angle will be noted as inconclusive. Figure 6 presents exemplary results that show the diversity of points chosen by the radial search after the 12 filter variations.

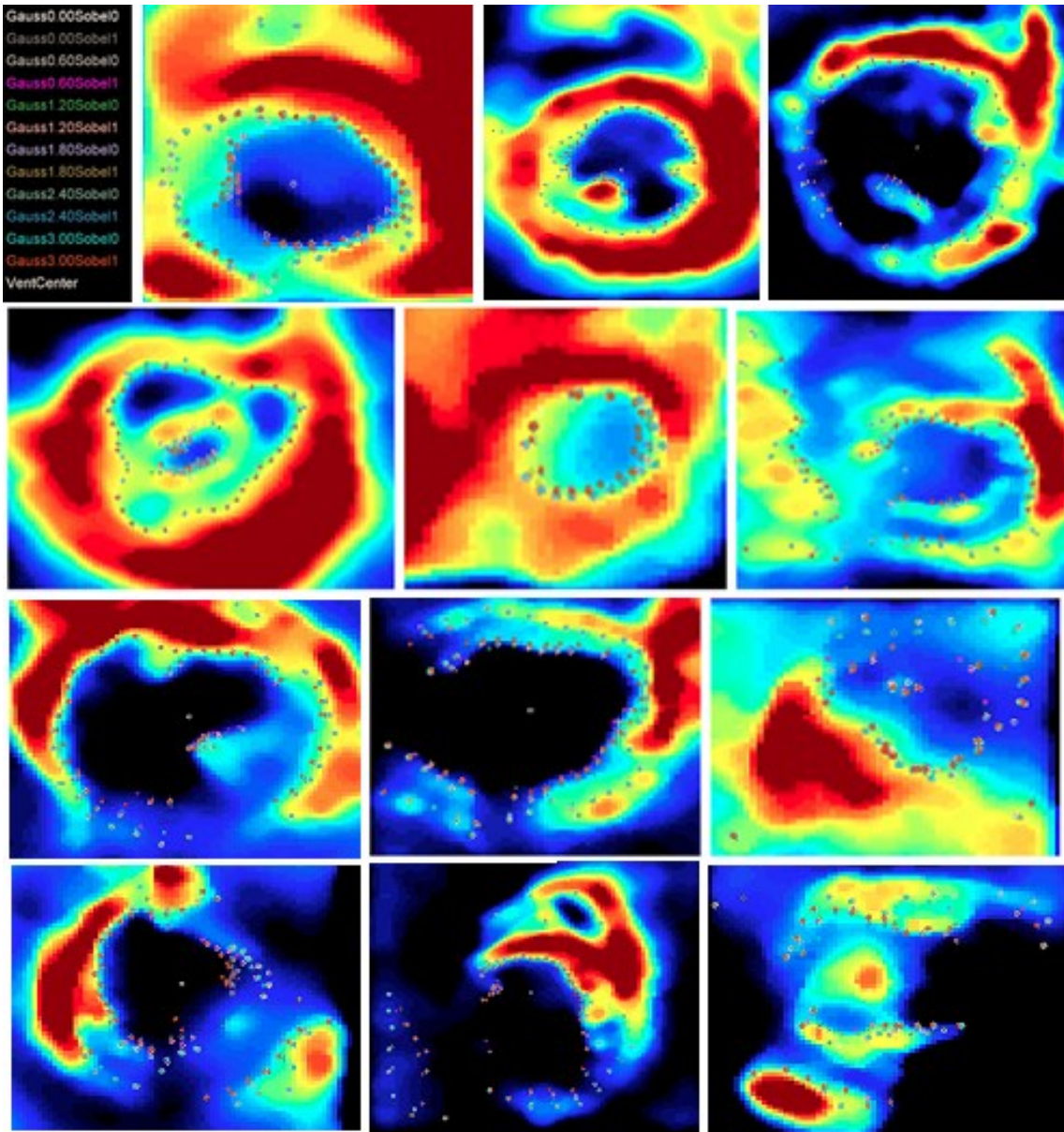


Figure 6: Example radial search and feature detection using all 12 filters to identify potential boundaries. Note how all filter variants are necessary to detect all the features ultimately required to draw the endocardial boundary. Any one filter on its own does not detect all the correct features of interest.

2.6 Ventricle Wall Census

For each angle of each short-axis slice, potentially 12 boundary estimates from the last stage now need to be reduced to a single census. The mean and variance of the radial distances between the previously found centroid and the boundary estimates are obtained for each angle. Outliers more than two standard deviations from the mean are discarded and the mean is recalculated. The number of boundary estimates for this angle that were both conclusive in the last stage and kept as non-outliers are assigned to be the “weight” for this angle. Outliers among the series of mean radii over all angles for the slice are then eliminated as per the protocol:

1. Any angles whose mean radii are zero will be reset if and only if less than half the angles have non-zero mean radii. They are reset to the series mean absent themselves and given unity weight.
2. Any angles whose mean radii are more than double the average of the series after (1) are reset to this average and half weight.
3. Any angles whose mean radii are more than 1.5 of the series average after (1) are assigned reduced weight by half.
4. Any angles whose mean radii are less than a fifth of the series average after (1) are reset to this average and half weight.

This protocol enforces correct declaration of the shallowest depth of the ventricle at the apical end without including tissue. It also ensures that unimportant artifacts in

the middle of the ventricle do not compromise the segmentation. Most importantly, the protocol eliminates points that may be far outside the ventricle or near the center due to a total absence of contrast of the true wall. In essence, the protocol attempts to force an elliptical shape where there is little reliability in border detection.

To fit a curve through the series of census radii, the robust LOWESS method is used again. The robust variant, as compared with the non-robust counterpart, offers additional protection from outliers. Both the linear and quadratic sliding window models worked well for a window size of approximately a third of the circumference. Also fitted are 4th-order and 5th-order polynomial curves to the radii series. The median of these four fits offered the best compromise by visual inspection of all slices. Select examples are shown in Figure 7.

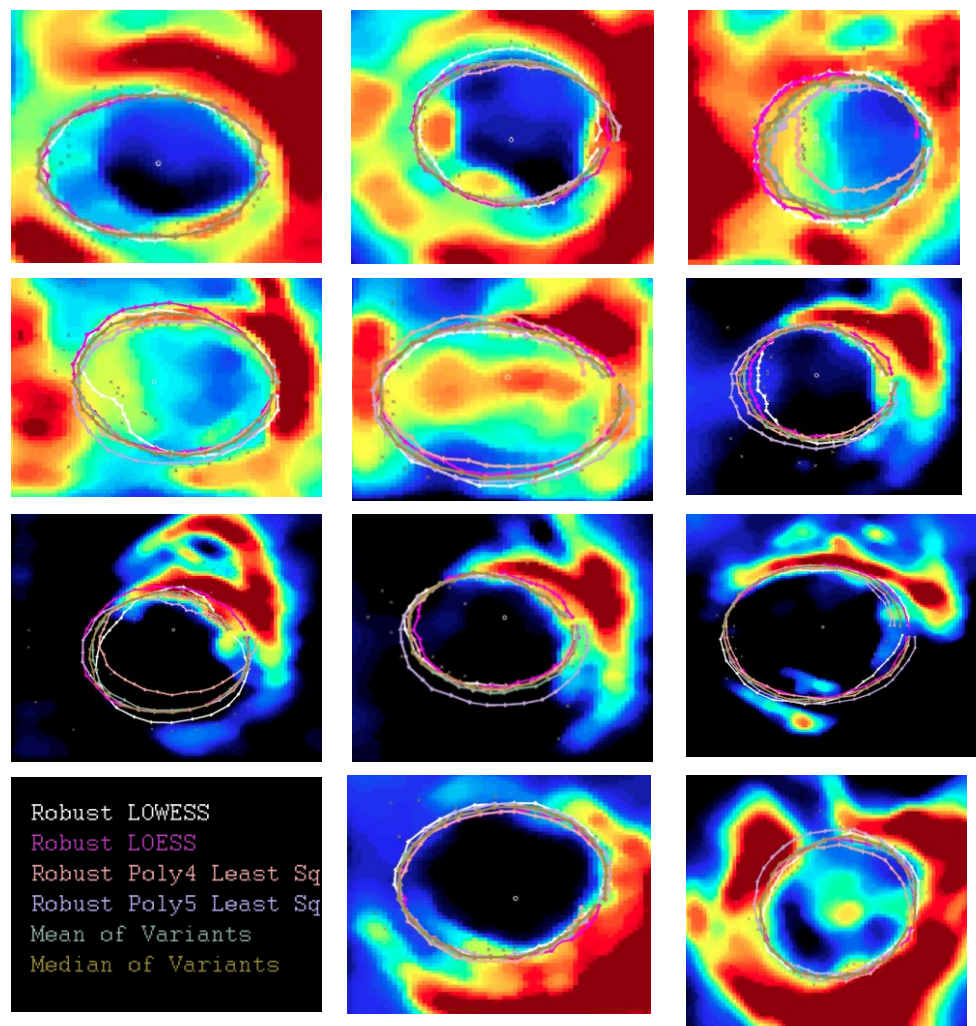


Figure 7: After the radial search features are merged to form census radii for each angle, outliers are removed from the radii series as per my protocol, and then curve-fitting through the radii series is performed. The algorithm proceeds with the medians of these curve fits.

2.7 Segmentation Optimization

Short-axis radial profiles obtained by the median fit are then fine-tuned to correct for minor errors. Each angle is processed independently on each slice on each frame. A maximal margin of correction of $\pm\Delta R_{MAX}$ is permitted for each radius. For the i -th angle of the j -th slice (depth j), the radius $r(\theta_i, j)$ is denoted r_0 and the objective function Obj is defined in (1):

$$\text{For } -\Delta R_{MAX} \leq \Delta r \leq \Delta R_{MAX}, \quad Obj(\Delta r) = \begin{cases} \sum_{r=R_U(\Delta r)}^{r=R_L(\Delta r)} \left(\begin{array}{l} \alpha(r) \cdot \left[\begin{array}{l} (slice(r, \theta_i, j) \geq LTL) \\ AND \\ (slice(r, \theta_i, j) \leq LTU) \end{array} \right] \\ + \beta(r) \cdot \left[\begin{array}{l} (slice(r, \theta_i, j) \geq HTL) \\ AND \\ (slice(r, \theta_i, j) \leq HTU) \end{array} \right] \end{array} \right), & |\Delta r| > 0 \\ 0, & |\Delta r| = 0 \end{cases} \quad (1)$$

Note the logic inside the square brackets in (1) is Boolean algebra. This significances of the constants are explained later. One also needs to define (2) and (3):

$$R_L(\Delta r) = \begin{cases} r_0 + 1 & , \Delta r \geq 1 \\ r_0 + \Delta r & , \Delta r \leq -1 \end{cases} \quad (2)$$

$$R_U(\Delta r) = \begin{cases} r_0 + \Delta r & , \Delta r \geq 1 \\ r_0 - 1 & , \Delta r \leq -1 \end{cases} \quad (3)$$

The simplest settings for $\alpha(r)$ and $\beta(r)$ that work at least as well as any others are given by (4) and (5):

$$\alpha(r) = \begin{cases} +1 & , \Delta r \geq 1 \\ -1 & , \Delta r \leq -1 \end{cases} \quad (4)$$

$$\beta(r) = \begin{cases} -1 & , \Delta r \geq 1 \\ +1 & , \Delta r \leq -1 \end{cases} \quad (5)$$

After computing the objective function for angle θ_i and depth j for the range of Δr in (1), which can be done in $O(\Delta R_{MAX})$ time if an accumulator method is used, one seeks to find the maximum of Obj and use this to reset $r(\theta_i, j)$ to the value $(r_0 + \Delta r')$, with $\Delta r'$ given by (6):

$$\Delta r' = \underset{\Delta r}{\operatorname{argmax}} \{Obj(\Delta r)\} \quad (6)$$

In the case of ties, the algorithm takes the Δr with the smallest magnitude. If there is still a tie, i.e. between a positive and negative Δr of the same magnitude, it leaves the radius unchanged because there is no way of knowing which direction of change is truly the better option.

Suitable settings for ΔR_{MAX} , LTL , LTU , HTL , and HTU were determined empirically using all datasets available. The key idea is to penalize the radial extension of the sector if it tries to grow through high-intensity backscatters, intensities between (HTL, HTU) , as it has then probably hit the wall or obstacles other than blood. The objective rewards radial shrinkage that eliminates encapsulation of these features. Similarly, the objective rewards extension where there is only fluid, and penalizes the loss of fluid area. Rewards are balanced against costs when one solves for the best radial tweak in each sector. It is critical to limit the ΔR_{MAX} setting, otherwise the optimizer may try to expand far out where there is little image intensity and, in doing so, outweigh the “cost” of going through the whole myocardium using sufficient “benefit” from low-intensity surrounding anatomy. For the same reason, it is critical to narrow the $(LTL,$

LTU) range so that one does not get “radial runaway” in the partial absence of wall contrast in poor images. I chose ΔR_{MAX} and intensity thresholds at values that emphasize radial shrinkage more than growth. The final settings offer considerable benefit according to visual inspection on all 7 test cases.

Next I apply a depth-wise filter to $r(\theta_i, j)$ through all depths j , carried out one angle θ_i at a time. For this, the LOWESS method was once again chosen with a moving window of 10% of the volume depth. By filtering each angle separately through all depths, one permits flexible shapes for the short-axis slices while enforcing smooth continuity of the ventricle depth-wise. The filter also corrects any major, isolated mistakes made by the optimizer. I show select examples of segmentation curve fits pre-optimization, post-optimization, and post-filter in Figure 8.

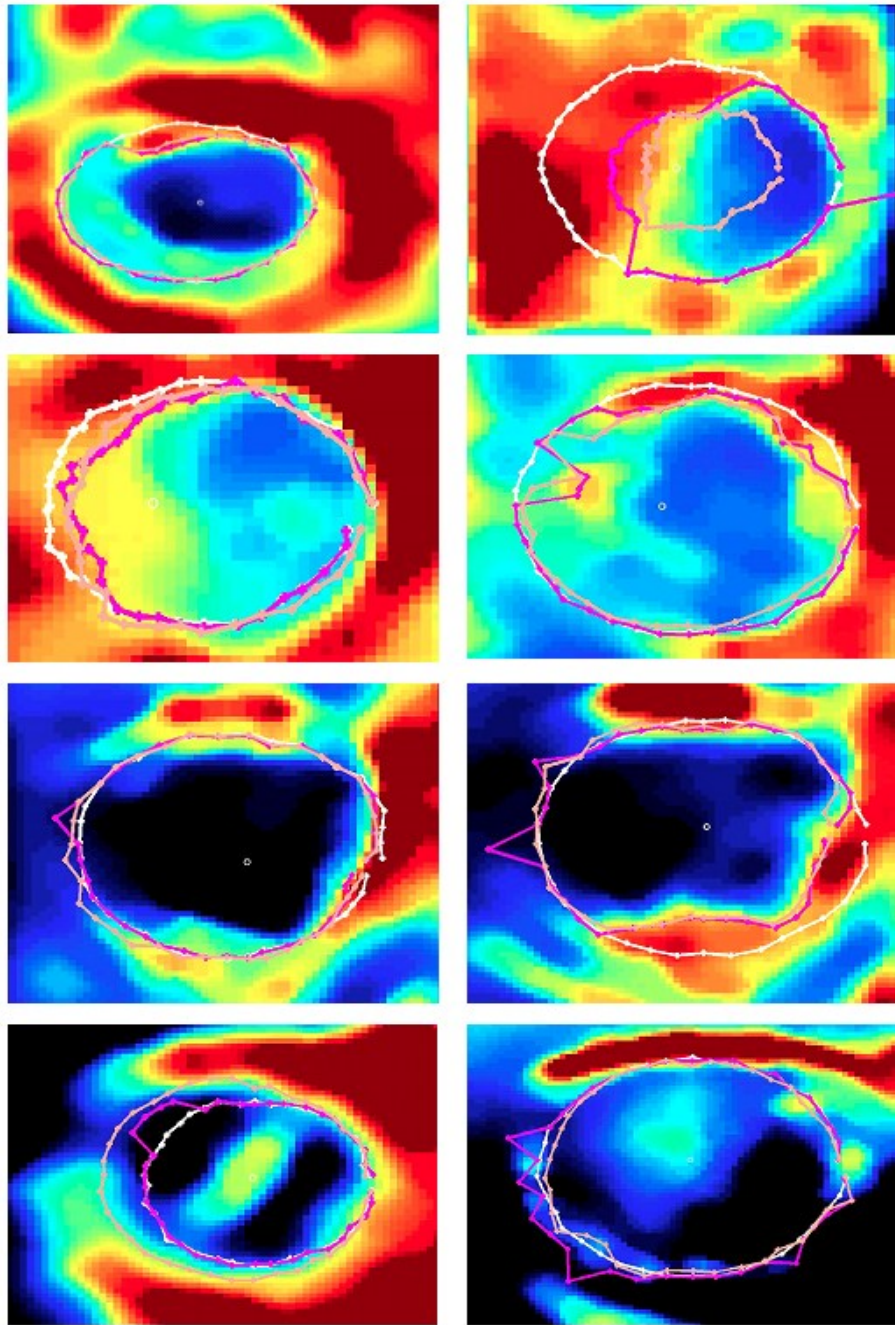


Figure 8: The median curve fit (white) is passed through the optimizer to produce another curve (purple). These selected examples show the optimizer adds small perturbations that improve the integrity of the segmentation. Finally, the radius of each angle is extracted from the optimized fit, and a 1D smoothing filter is used in the depth dimension on each angle to produce the final curve fit (peach).

2.8 Slice Area Calculation

With finalized $r(\theta_i, j)$ for all frames, the areas of the slices are trivially computed by considering the contributions of each sector. Each sector's radius is assumed constant as per a zero-order-hold interpolant. Other methods such as linear interpolation or polyphase upsampling had negligible effect on areas when at least 40 angles were used.

2.9 LV-LA Depth Delineation: Method 1

The last challenge of the segmentation is identifying a particular depth boundary between the LV and LA. First, I narrow a common search region that will be shared by all frames to between 50% and 80% of a physical depth reference. All standard apical images I have seen to date by trained sonographers suggest this is a generous range. Within this range, I identify the minimal slice area and use this to further trim the maximum depth of the search region. Then, starting from the shallowest depth near the apex of the LV, and using my algorithm's computed areas, I track the maximum area cumulatively encountered as I iterate through slices toward the base. As soon as I find an area that is 25% of the maximum I have seen so far, I cut the remaining depths out of the search region. The maximum will almost always be somewhere in the middle of the LV's depth span, and the small area will virtually always be outside the ventricle.

I next calculate the squared distance between the pre-filtered and post-filtered centroid point estimates. I found that this distance tends to show a major transient signal

when I hit the LV-LA boundary since there a small shift in the long-axis of the atrium when I leave the ventricle. The signal is more easily seen by taking a moving average window with size 5% of the depth. Even geometrical differences of the short axis slices above and below the mitral valve cause different long-axis estimates. I then get an estimate for the LV-LA delineation by seeking the maximal magnitude of this filtered signal within the previously established search region. This is repeated for all frames in the dataset and I take the median depth. Noting that the LV-LA boundary tends to never move more than 10% of the depth within the cardiac cycle, another generous number, I further refine the search region about this median. The refined search region offers greater consistency of my LV-LA boundary between frames by removing outliers that would greatly impact the volume curves. Finally, the search region is once more used to find a peak of the difference signal for each frame, and these are taken as the frame-wise LV-LA boundary depths. An example is shown in Figure 9.

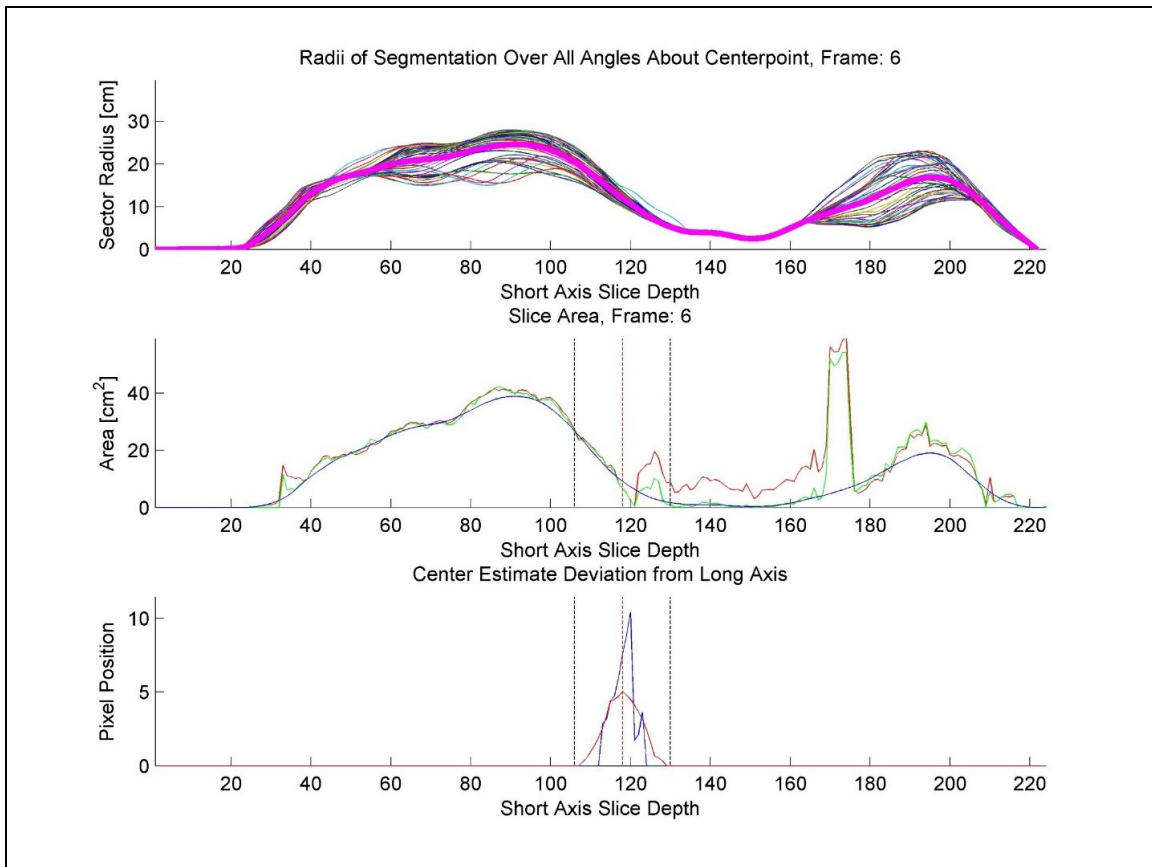


Figure 9: For one selected frame, the sector radii are shown along with the average radius (magenta) at every depth through the sample (top). The area of the ventricle at every depth, as segmented by the curve before the optimizer is used (red), post-optimization (green) and post-filtering (blue), is shown (middle). The final search region for the boundary between the atrium and ventricle is delimited to the same region for all frames, and the maximum of the moving average within this region of the deviation of the unfiltered ventricle center estimate from the long axis estimate is used to choose the separation depth (bottom).

2.10 LV-LA Depth Delineation: Method 2

The first method of LV-LA depth delineation (Section 2.9) is taken to be inconclusive if the shift of the long axis is insignificant. For these cases, a second method

was needed. This method attempts to estimate the depth of the mitral annulus as a function of frame. The method is based on a key insight discovered among a set of 70 training cases, without any exceptions, that the fibrous skeletal structure of the mitral annulus may be distinguished by a high-intensity backscatter signal. Searching for an approximate depth of the mitral annulus in the 4-chamber long axis view amounts to identifying this feature in the image. An example is shown in Figure 10.

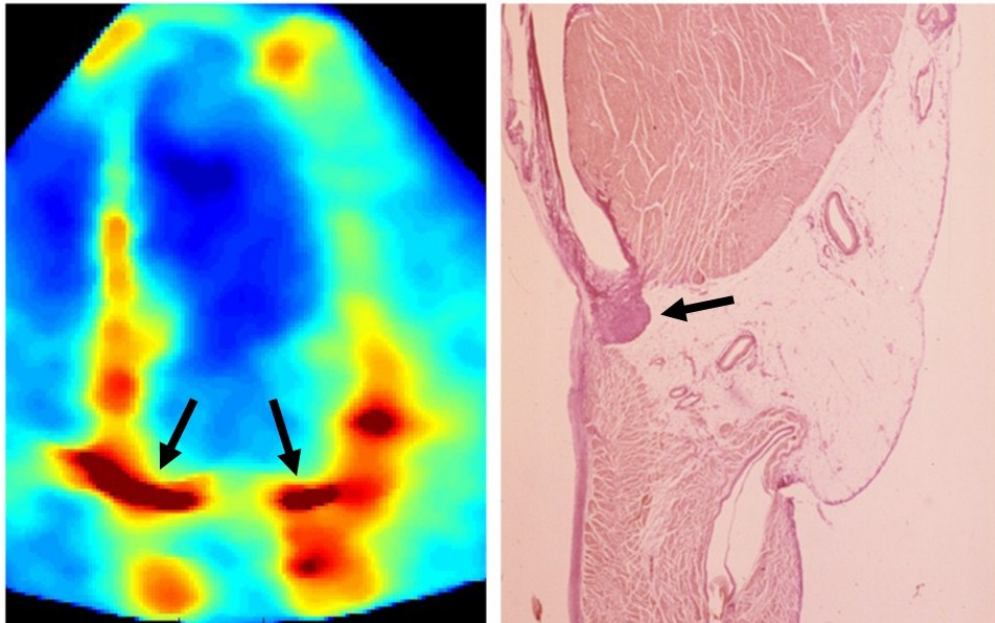


Figure 10. Histologically stained section (right) shows the fibrous skeletal feature of the mitral annulus, with correlation to the same feature in the 4-chamber image (left). Credit: Anderson, R. & Becker, A.: The Cardiac Skeleton. Gower Medical Publishing, Ltd. 1980.

To identify this feature, a range of 4-chamber long axis slices are taken within a range of +/- 10% of the height dimension (into and out of the page in Figure 10Error!

Reference source not found.) A filtering pipeline similar to that used on the short axis (Sections 2.2 through 2.6) is employed on these select long axis slices. Figure 11 shows intermediate results of this pipeline.

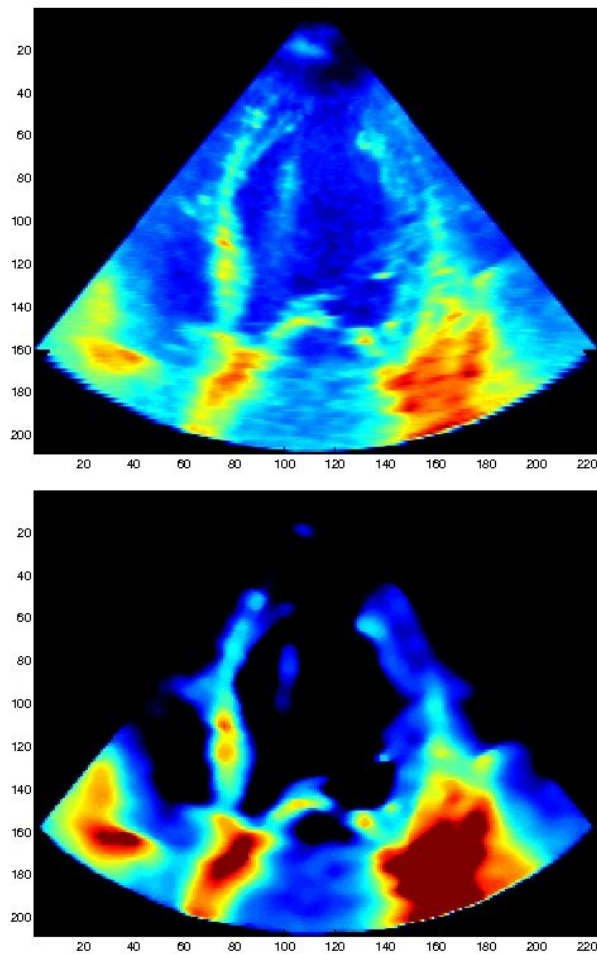


Figure 11. Long-Axis Filtering prior to feature detection for mitral annulus detection. Top: After a 1D Blackman impulse response is applied in the height dimension (face the page). Bottom: 2D filtering (same algorithm as Section 2.3) by contrast adjustment, Wiener filter, and order statistic filter.

Centroid estimation is applied on the long-axis using the filtered slices in a similar fashion as the short-axis (Section 2.4). However, the major difference is that the long-axis variant is intentionally non-converging. It was found the centroid tends to migrate toward the LA if permitted to converge. As long as the number of iterations was kept between 3 and 8, the results were acceptable on all 70 training cases. Hence, 5 iterations are used unless convergence is reached sooner. Using this centroid, feature detection is performed analogous to Section 2.5. An example of a long-axis centroid and feature detection is shown in Figure 12. An example of the iterative centroid locator is shown in Figure 13.

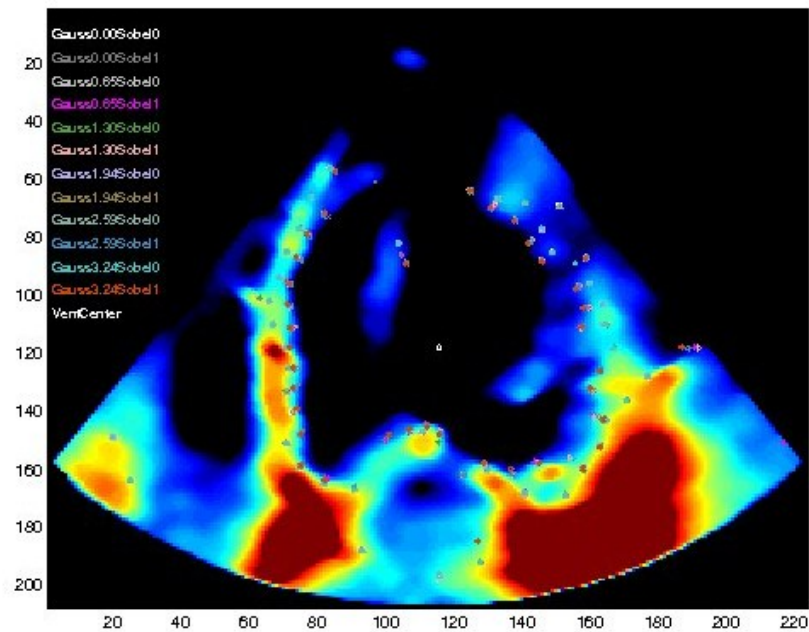


Figure 12. Feature detection method applied on the short axis is also used for the 4-chamber long-axis slices.

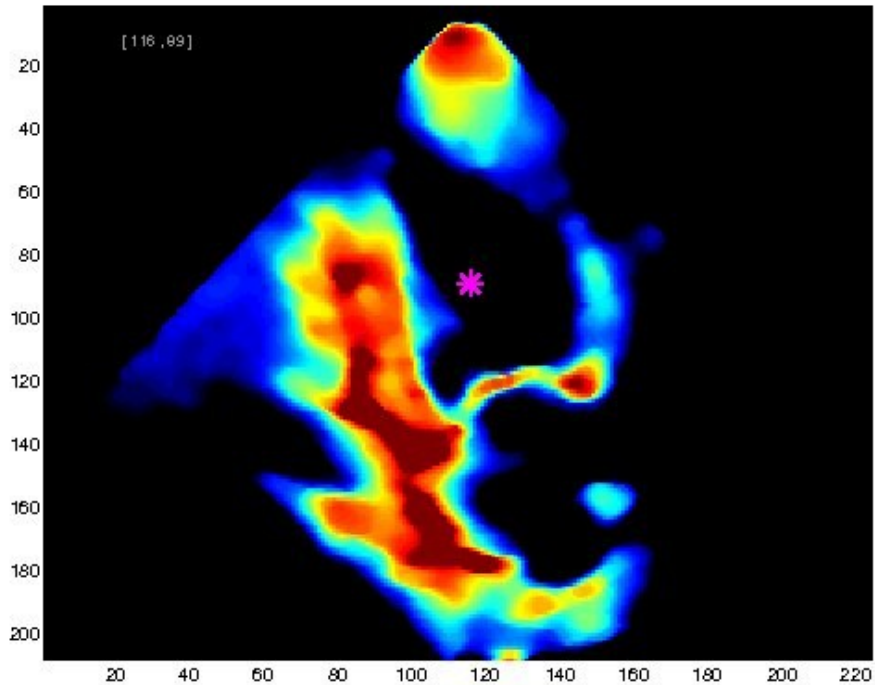


Figure 13. Example of centroid (marked by star) located on a column-wise slice (4-chamber, even though only 2 chambers may be clearly visible).

A mask is applied to select only the relevant points out of the feature detector. This mask selects only those points which occur on the high-intensity voxels, defined as being in the upper quartile of the filtered image intensity range. The median of each radial profile of points that successfully pass this mask is taken. Median points that are above the centroid estimate are removed. To isolate points that may be relevant to the mitral annulus, clusters typical of the high-contrast septal wall are detected and removed. A slope requirement between successive points is also used to enforce selection of points that likely lie on the annulus, noting that the boundaries of the

annulus tend to be orthogonal to the centroid chosen. After applying these filters, a typical set of annotations like Figure 14 is attained.

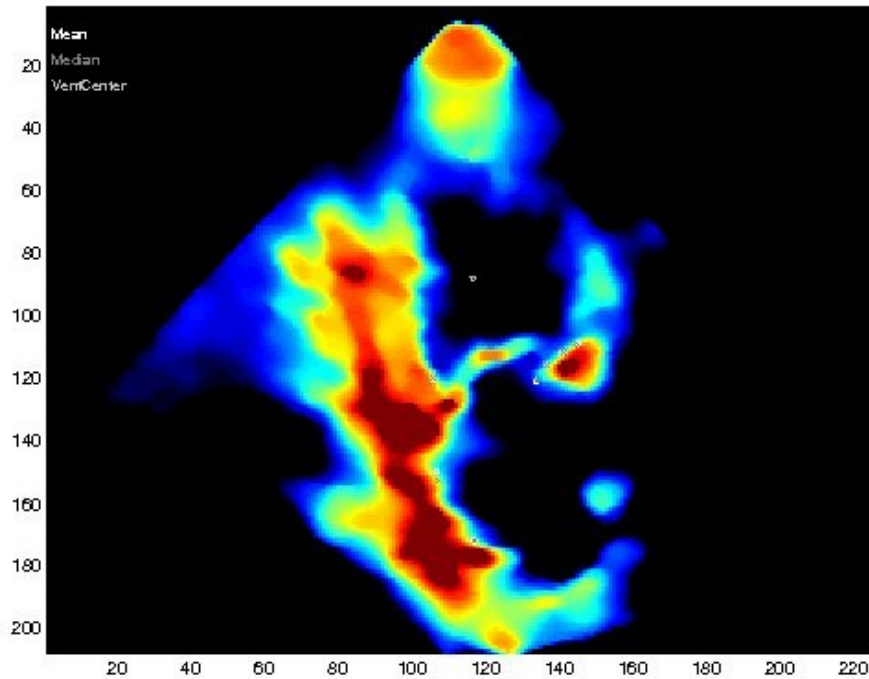


Figure 14. Points on the mitral annulus possibly identified after filtering the feature detection output on the long axis. The ventricle center is shown along with six points of interest.

The median depth of these points is determined. Then these medians are used to find yet another median through all height slices for each frame. Thereafter, an optimization step is used to locally shift the median depth such that the sum of intensities over the line is maximized within an adaptive window. This optimizer runs for each column-slice independently. The median depth is again taken through all of the post-optimization depths of all the studied columns for each frame. Finally, the frame-

series of depths undergoes robust moving window linear regression to smoothen the series and remove outliers. A typical result of these steps is shown in Figure 15.

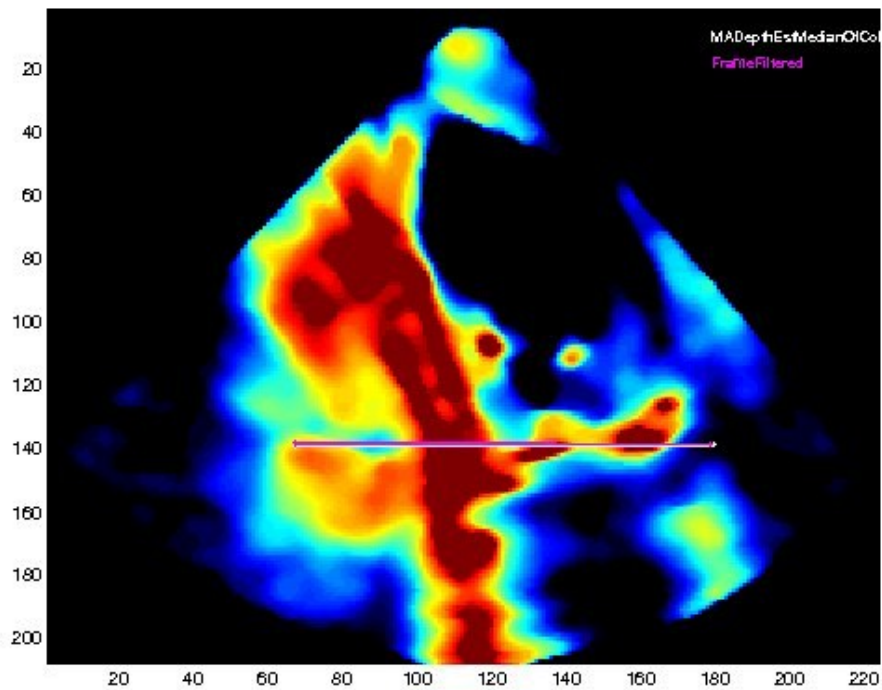


Figure 15. Example of the LV-LA delineation being approximated by an estimated determination of the depth of the mitral annulus, for a particular frame, shown on top of the 4-chamber view.

2.11 LV Volume & EF Calculation

Summing slice areas in their physical units up to the detected base boundaries and multiplying through by slice thickness yields my LV volume. Repeating this for every frame allows us to characterize volume over time and calculate EF, as shown in Figure 16. Shells of the volumes are constructed by tracing through the short-axis

circumference over all depths on top of zero-initialized volumes. To produce smooth-looking shells, I upsample the radial profiles. I can also visualize the segmentation in context, as in Figure 17, which shows the luminal shell 3D-rendered on top of the anatomy. A number of 2D planar cuts highlighting the segmentation are given in Figures Figure 18 and Figure 19.

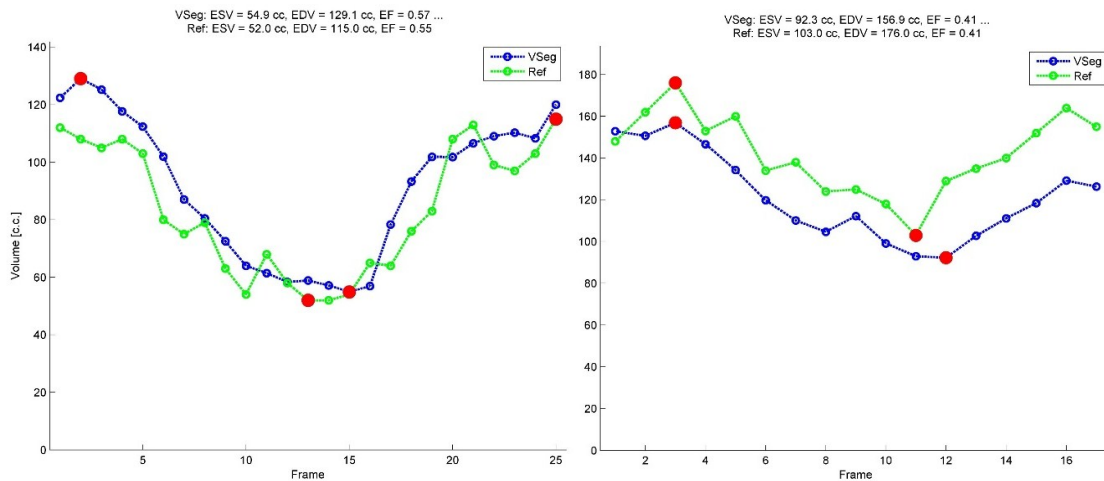


Figure 16: Volume time series for two cases. The output of the algorithm is compared to a manual segmentation.

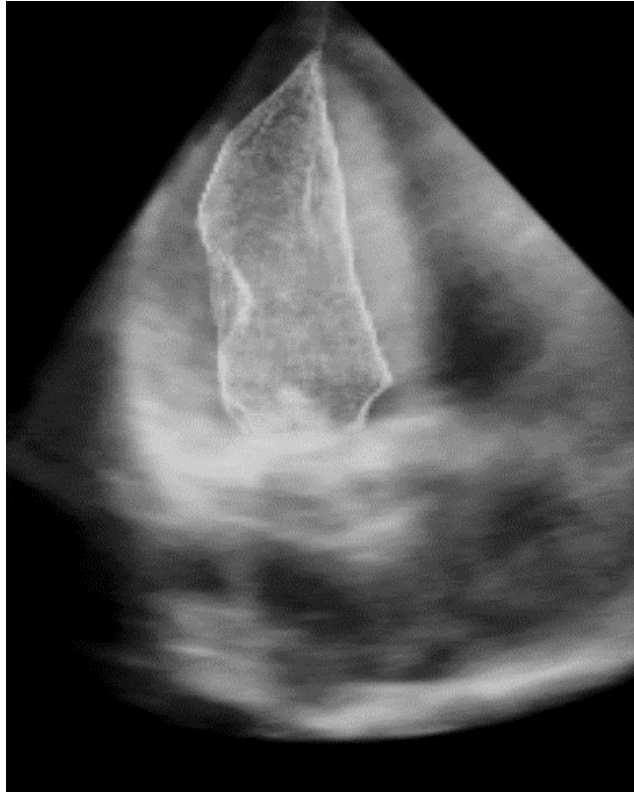


Figure 17: The boundary of the segmented lumen can be viewed in context with the rest of the anatomy using a 3D volume rendering. This methodology was part of the verification process, and trivially implemented using the segmented volume.

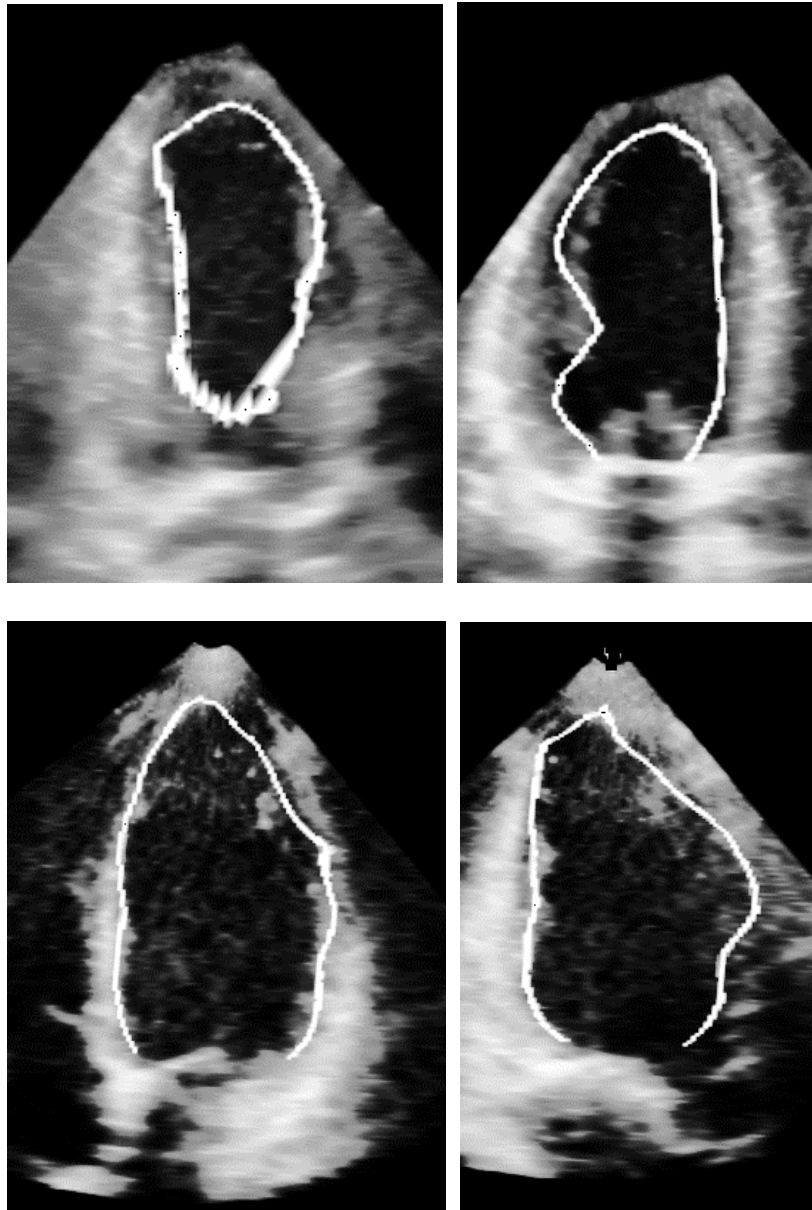


Figure 18: (Top) Two orthogonal planar selections of a dataset's EDV with segmentation boundaries overlaid. (Bottom) Another two orthogonal planar selections of a different dataset also in EDV with segmentation boundaries shown.

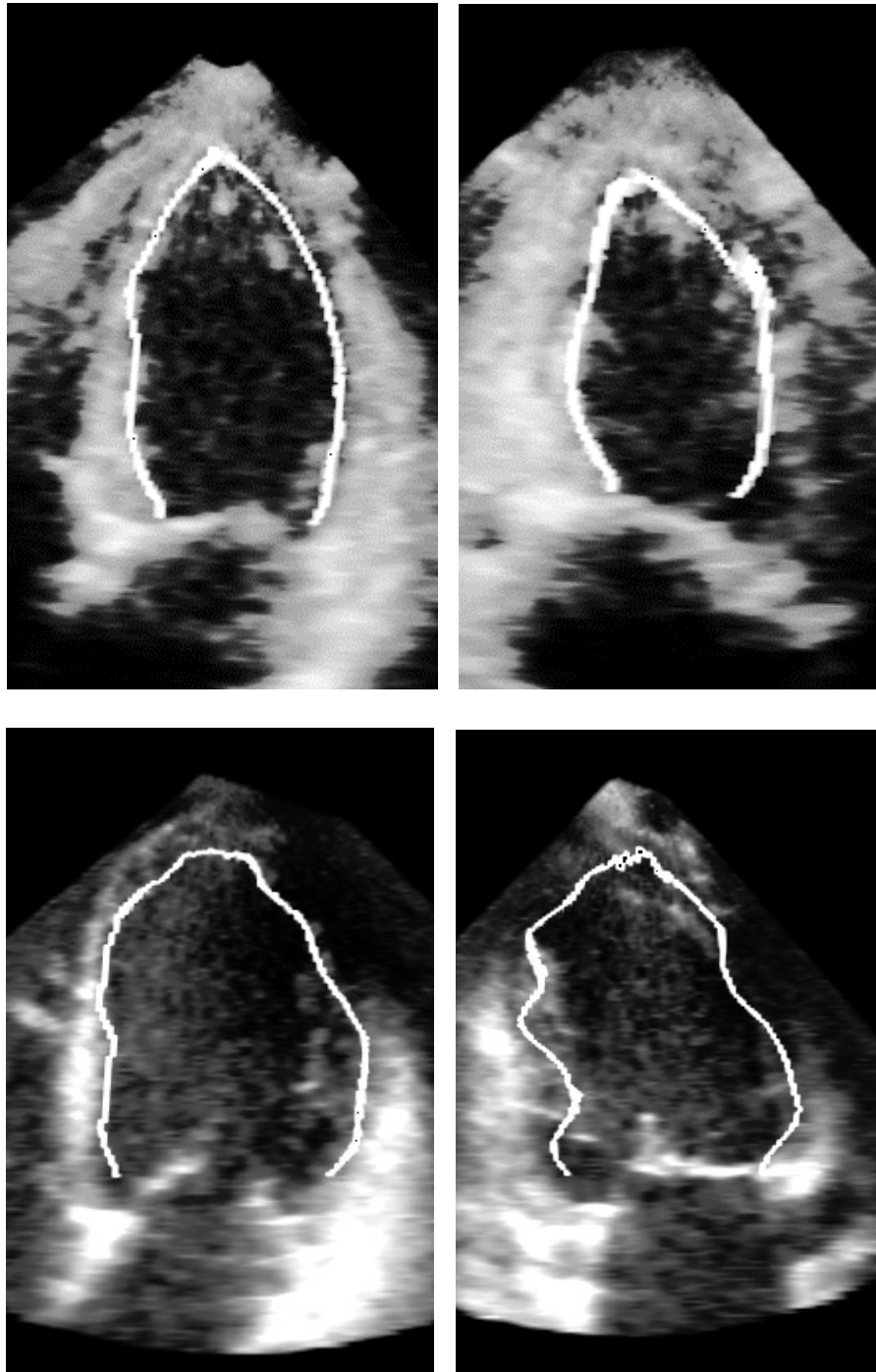


Figure 19: (Top) Two orthogonal planar selections of a dataset's ESV with segmentation boundaries overlaid. (Bottom) Another two orthogonal planar selections of a different dataset also in ESV with segmentation boundaries shown.

3. Results & Pre-Clinical Validation

A set of 70 patients undergoing treatment at Duke University Hospital doubled as both a training set for iterative algorithm development and a pre-clinical validation set. Using this training set, the method was validated against both manual segmentations, as obtained using the Duke 4DViz software tool, and the semi-automatic Philips QLab-3DQ volume quantification tool. The latter software was previously validated against cardiac MR in a multi-center study [10], and serves as a gold standard for the purpose of validating my method.

3.1 Comparison to Manual Segmentation

The first 7 of the 70 cases underwent manual lumen-endocardium border segmentation using the Duke 4DViz software by 2 or 3 operators. The algorithm was also run on these cases. Volumetric models were output for visualization of the segmentations using both softwares. All manual segmentations were found to be sensible. After considering overlap or gaps between the segmentation boundaries and the blood-endocardial boundary, my clinical collaborators concluded that they were accurate without any major errors.

I then studied the volume time series of each patient out of the algorithm and compared the trends and numbers to those of the manual segmentations. In Figure 16 I showed an example. The cardiac cycles are detected and mostly synchronous between

the methods. Using the minimum and maximum volumes of each time series, EF was calculated and compared for the manual and automatic methods, reported in Table 1. The algorithm predicts EF for 6 of the 7 patients with ± 0.09 units of difference from the manual, well within common clinical tolerance [9]. One patient showed discordance of 0.15 units, and it remains unclear which method offers the more reliable EF for this case.

Table 1. Comparison of ejection fractions for 7 training cases as determined by the algorithm to those obtained from manually segmented volumes.

Case	Ejection Fraction	
	<i>Algorithm (Automatic)</i>	<i>Manual Segmentation (Average)</i>
1	30 %	30 %
2	56 %	55 %
3	44 %	41 %
4	14 %	29 %
5	23 %	22 %
6	23 %	26 %
7	40 %	31 %

Visual comparison of the automatic and manual segmentations overlaid on the original volumes, viewed both with volume rendering and cut planes, was done using software developed by my group. The visualizations suggested that the automatic method may offer more reliable segmentations than those acquired manually in any reasonable timeframe, possibly explaining the larger discrepancy of EF observed for two

of the seven patients. The luminal shell of each automatic segmentation visually appeared plausible when rendered within anatomical context, as in Figure 17.

Another interesting observation is that while EF between different raters and methods may be concordant, absolute volumes between them tend to be systematically offset by a multiplicative factor. It is mathematically the case that if this factor stays relatively constant between frames, there should be little difference in EF between segmentation methods applied to the same dataset. This may explain why ejection fractions are sometimes remarkably similar even though the ESV and EDVs may vary by tens of milliliters between methods. This may be due to different opinions on where to draw the LV-LA boundary, the coarseness of the ventricle wall, confusion of the trabeculae as the endocardial boundary, and inclusion or exclusion of the vicinity of the papillary muscles.

3.2 Comparison to QLab-3DQ

A subset of 13 cases selectively chosen because of their reasonable image quality were analyzed using the semi-automatic Philips QLab-3DQ software. One or two operators used the software to produce ejection fractions. The first operator was Wald, who investigated all cases and recorded the mean of 3 trials per case. The second operator was a random sonographer, possibly different between cases. The results are summarized in Table 2. The average operator data from 11 cases showed clinical

concordance (± 0.10) with the algorithm. The two remaining cases showed clinically unacceptable difference (defined as greater than ± 0.10) yet it is unclear which segmentation offers the more reliable estimate of EF. Most of the difference against QLab-3DQ was attributed to differential mitral annulus elucidation or the exclusion of suspected endocardium in the luminal space. Another important observation is that, as with the manual segmentation comparison, absolute volumes produced by the algorithm tend to be systematically off-set by a multiplicative factor between the two methods. As was true with manual segmentation, this factor is again relatively constant between frames, cancelling out for minimal effect on EF.

Table 2: Comparison of ejection fraction as determined by algorithm to that determined by QLab-3QD as used by an experienced sonographer and/or Wald

Case	EF: AutoV [%]	EF: Duke CDU Sonographer QLab [%]	EF: Wald QLab [%]
8	32	40	40
9	53	42	43
10	45	59	65
11	23	14	27
12	23	29	28
13	43	---	47
14	14	21	20
15	17	25	20
16	56	62	58
17	50	31	48
18	56	---	56
19	33	35	34
20	30	28	30

The remaining 50 cases of the training set were chronologically acquired retrospectively. These cases represent a mix of good, mediocre, and poor image quality. Some sonographers acquired data for all-comers. Other sonographers only collected data for a patient if they determined that the patient produces images with quality suitable for 3D quantitative analysis. Hence, the bias toward good quality images in this subset of 50 cases is unknown, but certainly lower than the subset of 13 cases in Table 2. In this subset, 62% of cases show concordance (± 0.10), and 23% show unacceptable difference between 0.11 and 0.20 units of EF. A further 15% of cases show discordance beyond 0.20 units. While investigating the larger frequency of algorithm discordance in this subset, it was determined that image quality may be a major determinant of accuracy. The algorithm performs well on good images almost every time. Furthermore, the automatic segmentation may be preferable to QLab on some datasets, highlighting the potential pitfall of using the Philips implementation as a gold standard for this study. Poor quality will often predict discordance. Image quality is generally dependent on the individual, such as adipose located between the transducer and apex.

By far, the largest challenge to my fully automatic method was robust detection of the mitral annulus. As QLab requires at least 4 points on the annulus to be specified by the user, the Philips system has an inherent mechanism to overcome this issue. My method for annulus detection, which depends on the highly-backscattering fibrous skeleton, generally works well. However, failure to correctly pick the annulus typically

leads to an erroneous volume-time curve, and often an incorrect EF. A second common error explaining the discordant cases is that excessive missing wall contrast was most likely to occur in diastole and least likely in systole. EF was overestimated in these cases. In some instances, artifacts of unknown origin, perhaps from tumors or detached papillary muscles, can mask key features such as the apex or be misinterpreted by the algorithm as endocardium.

4. Clinical Trial Design

To validate my algorithm, I have initiated a retrospective clinical trial (Duke CDU IRB DELD-050) using 3D apical echocardiographs collected at Duke University Hospital. Some sonographers were uninstructed on the trial and proceeded to choose which patients qualified for apical echoes based on image quality, need, or other factors. Other sonographers were explicitly instructed to collect data on all-comers regardless of image quality. All datasets collected at the Duke Cardiac Diagnostic Unit (CDU) with 3D apical echoes were recruited into the study based on a chronological start and end point in order to reach N=60. Like the last 50 cases of the training set, the 60 cases of the validation set will be a mix of previously acquired all-comers and good quality cases (sonographer voluntarily collected). Three senior sonographers with the Duke CDU have been recruited into the study to investigate these cases without collaboration. Each sonographer will independently use the Philips QLab-3DQ software with minimal instruction to produce EDV, ESV, and EF numbers for each case. They will be blinded to any such existing analysis for any case. They will subsequently run my algorithm on the same datasets. The trial will seek to compare the results of the two methods. The sonographers will also investigate plausibility of EF and volumes out of the two segmentation methods, and rate image quality for each case on a three-level scale. I originally hypothesized that my automatic method can provide results highly comparable to those of the sonographers using QLab-3DQ (within acceptable clinical

tolerance), without requiring any sonographer input other than runtime initiation. However, based on my pre-clinical validation results in Section 3.2, a more suitable, revised hypothesis would be that the algorithm can automatically reproduce QLab-3DQ LVEF for a major subset of patients with sufficient image quality, where quality is measured by a metric yet to be determined. This concept will need to be further explored in the follow-up clinical study once the team of sonographers have produced reliable EFs.

A second retrospective clinical trial I have initiated will compare my algorithm with LV volumes from cardiac MRI segmentations as a gold standard. The cases for this trial will be new and separate from those of previous datasets in the study. The design of this trial is forthcoming at this time.

5. Conclusion

In this paper, I proposed a novel algorithm for computing LV volumes through the cardiac cycle and ejection fraction using apically-acquired 3D echoes. The algorithm adopts a number of feature detection methods to the 3D echocardiography modality. The algorithm pipeline was presented, along with design justifications. Pre-clinical validation was performed by both manual segmentation and comparison to Philips QLab-3DQ. These results indicate that my method has promise to automate volume quantification otherwise producible using manual or semi-automatic methods. A clinical trial has been initiated to further validate the method and begin the translational process.

At present, the algorithm has been prototyped in Matlab. Though it is not currently considered real-time by any reasonable measure, prior experience indicates that the prototype may be a prime candidate for software and/or hardware acceleration. Hence, the method has promise to provide instant clinical point-of-care volume quantification with further development.

References

- [1] Paragios, N. (2003). A level set approach for shape-driven segmentation and tracking of the left ventricle. *IEEE Trans. Med. Imag.*, 22(6): 773-776.
- [2] Bosch, J. G., et al. (2002). Automatic segmentation of echocardiographic sequences by active appearance motion models. *IEEE Trans. Med. Imag.*, 21(11): 1374-1383.
- [3] Angelini, E. D., et al. (2005). Segmentation of real-time three-dimensional ultrasound for quantification of ventricular function: A clinical study on right and left ventricles. *Ultrasound in Medicine and Biology*, 31(9): 1143-1158.
- [4] Yezzi, A., Jr., et al. (1997). A geometric snake model for segmentation of medical imagery. *IEEE Trans. Med. Imag.*, 16(2): 199-209.
- [5] Van Stralen, M., et al. (2007). Automatic Segmentation of the Left Ventricle in 3D Echocardiography Using Active Appearance Models. *IEEE Ultrasonics Symposium, 2007*: 1480-1483.
- [6] Gerard, O., et al. (2002). Efficient model-based quantification of left ventricular function in 3-D echocardiography. *IEEE Trans. Med. Imag.*, 21(9): 1059-1068.
- [7] Lim, J. S. (1990). *Two-Dimensional Signal and Image Processing*, Englewood Cliffs, NJ, Prentice Hall, p. 548.
- [8] Cleveland, W. S. (1979). Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association*, 74 (368): 829-836.
- [9] Van Royen, N., et al. (1996). Comparison and reproducibility of visual echocardiographic and quantitative radionuclide left ventricular ejection fractions. *Amer. Jour. Card.*, 77(10): 843-850.
- [10] Mor-Avi V., et al. (2008). Real-Time 3-Dimensional Echocardiographic Quantification of Left Ventricular Volumes. *JACC: Cardiovascular Imaging*, 1(4): 413-423.