Ideal A-Scan Determination for Vibration Measurements within the Organ the Corti

Brian Frost
Electrical Engineering PhD Candidate
Columbia University
Email: bf2458@columbia.edu

***Abstract*—Spectral domain optical coherence tomography (SD-OCT) is commonly used in cochlear mechanics research. The depth images provided by SD-OCT allow for multiple cellular layers in the cochlea's Organ of Corti (OoC) to be distinguished, and sub-nanometer vibrations in these layers can be detected using the technique of spectral domain phase microscopy. For a given animal preparation, it is often difficult to obtain a "good" location for vibration measurements due to variations between animals. We present here a method to quickly determine the best location for measurements to be taken given a volumetric scan of the OoC through the round window, which can remove a time-consuming aspect of the set-up phase for vibration measurements almost entirely. We present also MATLAB code for this method implemented on a volumetric scan of an *ex vivo* gerbil cochlea.**

I. Introduction

At Columbia's Fowler Memorial Laboratory, we use SD-OCT and phase microscopy to measure vibrations at multiple positions in the OoC in gerbils and guinea pigs. After animals are anesthetized and surgery is performed, significant time is spent attempting to obtain a satisfactory position at which to take measurements. Time is precious in experiments of this nature as the animals cannot remain anesthetized forever, and so often suboptimal measurement locations are chosen out of expedience. Similarly, in cases where an experiment is ongoing and a condition changes (the animal's head is moved, the experimenter accidentally moves the OCT machine), a position must be re-chosen in even tighter time constraints.

Cochlear vibration measurements with SD-OCT are made along a single axial path, and the signal-to-noise ratio (SNR) of the vibration measurements is determined by the SNR of the A-Scan along that path. We are usually interested in measuring at least two locations within the cochlea: the basilar membrane (BM) and the outer hair cell (OHC) region. Using B-Scans or volumetric scans, these regions can be distinguished by a small fluid-filled space in the OoC, and an A-Scan which includes both regions may be chosen. Figure 1 shows an example of a "good" A-Scan from a gerbil cochlea, located by looking at the B-Scan on the left in an *in vivo* gerbil cochlea. This method, unfortunately, requires a high-SNR B-Scan to first be manually located by the experimenter, and this is rarely an easy task.

Here, we present a method that reduces the time significantly for finding a set of valid A-Scan locations at the beginning of an experiment of this type. The inputs to the algorithm are a volumetric OCT scan and an axis along which to consider it as a "stack of B-Scans". These stacked B-scans are then segmented using the assumed anatomy of the mammal's cochlea, and the OHC region is localized. Then, the highest-SNR A-Scan location in each B-Scan is presented to the user, and the best of these locations is presented to the user as well.

II. The Algorithm

A. The Input

The algorithm consists of a number of operations performed on each B-Scan, and the "stack" nature of the volumetric input is only considered in the very last step in which the "best points" for each B-Scan are compared to one another. That is to say that the input is interpreted as a group of independent two-dimensional eight-bit grayscale images of any size, and then the outputs of the algorithm for each image are compared and presented to the experimenter. The rest of the algorithm will be presented as being applied to a single B-Scan, with the understanding that the same step would be applied to each image in the stack. An example of a B-Scan in our sample volume appears in Fiure 2.

B. Median Filtering

SD-OCT images suffer from salt-and-pepper noise, as can very clearly be seen in Figure 2. The canonical solution to salt-and-pepper noise is the median filter, as it removes outlier pixels at both low and high intensities. However, the median filter does incur some smoothing, which we want to avoid in areas with fine features. We wish, most specifically, to maintain the ability to see the gap in the OoC, which is about 30-50 pixels at its widest in our images. We find that a 5-by-5 median filter is sufficient for removing salt-and-pepper noise without smoothing out these important features. The result of this step can be seen in Figure 3.

C. Gaussian Filtering

Our goal is to eventually segment the image, and to do so we will use edge detection. The use of a Gaussian filter is known to facilitate edge detection (see Gonzalez and Woods), but has a smoothing property that once again could close the gap we are interested in maintaining. We choose to use a 5-by-5 Gaussian filter with standard deviation 2. This size and standard deviation come from the values used in the Fowler Lab when presenting B-Scans in presentations and

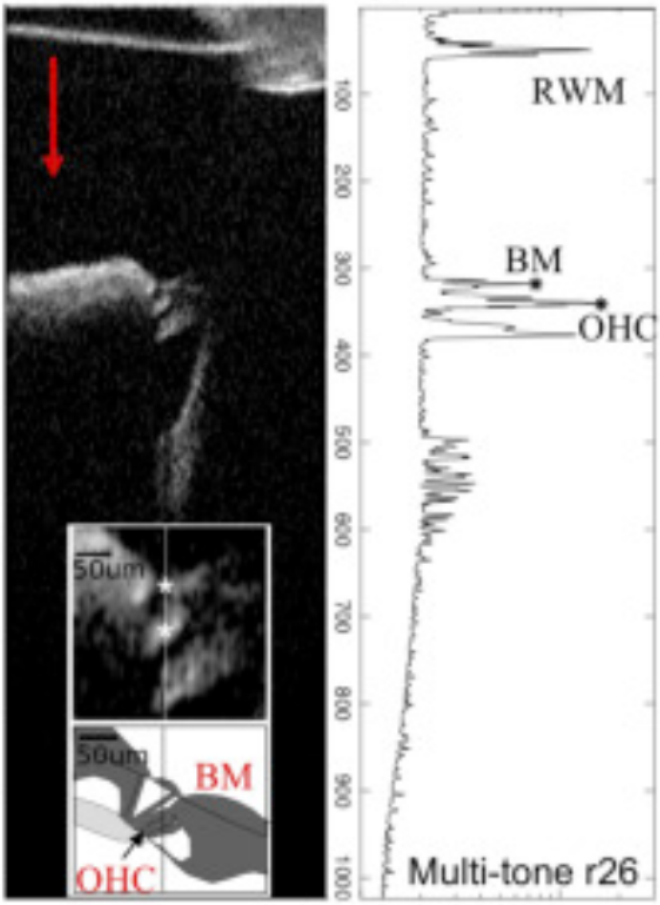

Fig. 1: An example of a "good" A-Scan, and the B-Scan from which it was taken, as well as the corresponding anatomy of the cochlea. The images are from an *in vivo* gerbil experiment at the Fowler Memorial Lab at Columbia University. The space between the BM and the OHC region yeilds distinct peaks separated by about 50 microns. RWM refers to the round window membrane; a membrane at the base of the cochlea through which the images are taken.

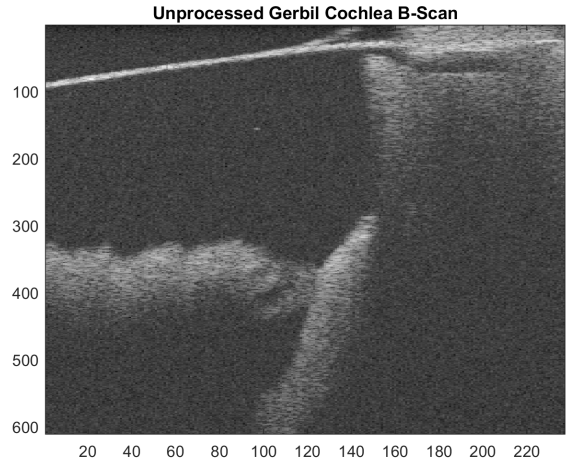

Fig. 2: An unprocessed B-Scan from a volumetric scan of an excised gerbil cochlea taken through the round window. Although there is noise, the BM, gap and OoC can be clearly made out around the 100-pixel x position. The processing of this image will be followed through the course of this document.

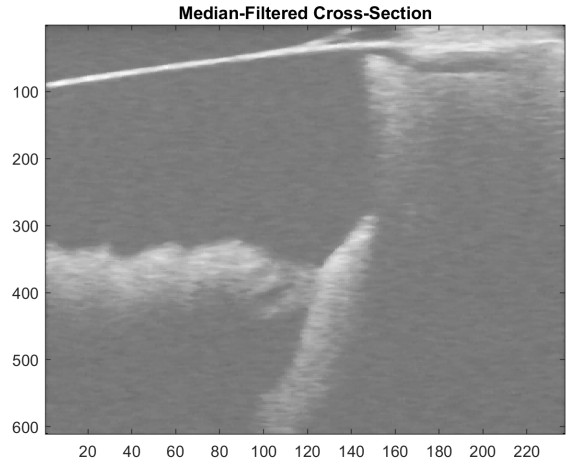

Fig. 3: Output of the median filter step of the algorithm. The effects of this filter are quite clear, as the features are smoothed out and the salt-and-pepper noise is reduced.

publications. It has been determined that it will not close gaps and will sufficiently smooth edges in a variety of images of the cochlea over extensive trial and error. The Gaussian filter must be applied after the median filter, so that it does not dilate the salt and pepper features. The output of this step of the algorithm is shown in Figure 4.

D. Thresholding and Contrast Enchancement

As a final step to facilitate edge detection, we maximize the contrast between foreground and background features. The fluid in the cochlea appears dark and counts as background, as do bone features which have become dark through obfuscation by shallower bone. We make these features totally black (intensity value 0) by mapping all intensity values under a threshold T to 0. Conversely, we map all intensity values above T to be totally white (intensity value 255). In the resulting image, an "edge" is very easily defined as a transition pixel – one which borders both a black and a white pixel. This is equivalent, in this case, to the Laplacian being nonzero. We choose a threshold of 40 for our images, but it will depend on the histogram of the image. The output of this step of the algorithm can be seen in Figure 5.

E. Edge Detection

For edge detection, we can use any number of methods as our images' edges are so well defined. We choose to use the built-in MATLAB function "edge" with the "canny" algorithm, as this method naturally connects edges. We need

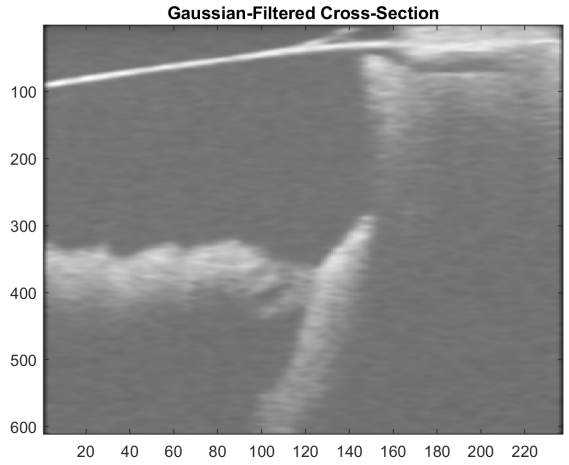

Fig. 4: Output of the Gaussian filter step of the algorithm. The effects of this filter are quite clear, as the features are smoothed out far more. The important gap is still visible, however.

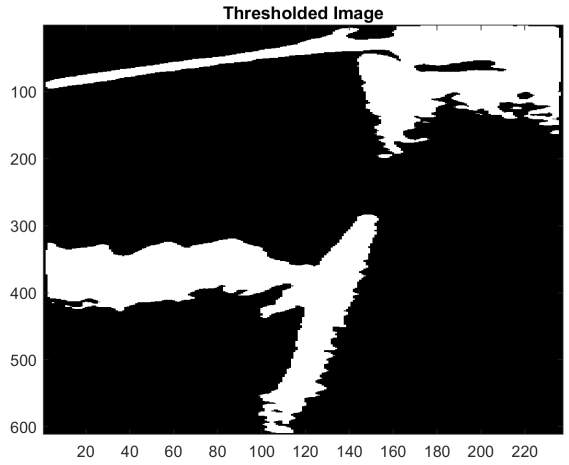

Fig. 5: Output of the thresholding step of the algorithm, in which all bone and tissue is represented in white. The important gap remains visible.

connected countours defining our region of interest (this will make sense shortly), so this method is particularly useful for our application. The results of this step are shown in Figure 6, where our region of interest an the associated gap are even more clear tothe observer than in our original image. It is also interesting that the BM and OoC (the horizontal feature) has been connected to the bony outer wall (the vertical feature) seamlessly. These objects are anatomically quite different, but the nature of OCT is such that they are indistinguishable due to their similar reflectivities. As intelligent observers, we can project the known anatomy onto the OCT image and interpret what we see, but this is difficult to make general enough to include in the algorithm. A solution to this problem is applied later in the algorithm.

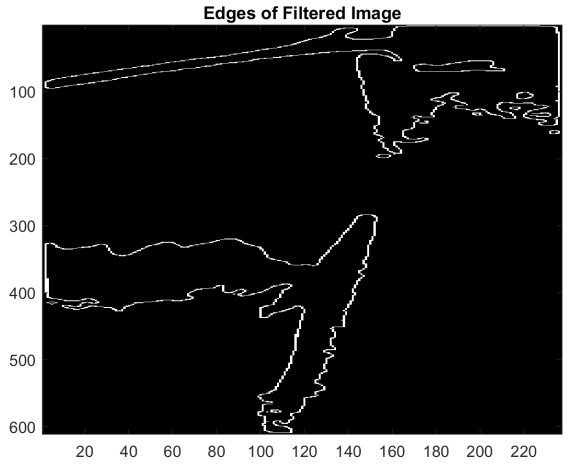

Fig. 6: The output of the edge detection step of our algorithm.

F. Primary Segmentation

To isolate the region containing the OHC region and BM, we first look to remove all of the other extraneous large features in the image. For example, the region at the top of the image containing the round window membrane and a chunk of bone is of no interest for our purposes. To do so, we realize that the BM and OHC region are contained in some large contour, which includes surrounding bone. This is, in all images of interest, the largest connected contour in the image. To isolate this contour, we use the MATLAB function "bwconncomp" on the edge image. This function returns a cell array containing arrays of indices corresponding to connected white components in a black-and-white image. For the image above, we show the four largest of these components in Figure **??**. We always take the largest contour, as it is the one which contains the portion we are interested in every tested series of images. We remove all information outside of these contours by first filling the contour using "imfill". To fill the contour, you must first dilate the contour (using "imdilate" with a simple radius-2 disk structural element) slightly so as to thicken the edge. Once the contour is filled, you have a mask for your original image, outside of which you can delete all irrelevant content. We apply this mask to the edge image, and are left with a smaller superset of our region of interest.

G. Secondary Segmentation

As mentioned before, we are confronted with the issue that our region of interest is attached to bone which cannot be distinguished clearly. We now rely on the particular anatomy of our problem - our region of interest is distinguished by the gap we have made enormous effort to maintain. We refer around the Fowler Lab to good A-Scans as having "the three peaks", which is easily quantified in an edge image. As such, we first zero out every vertical line (or potential A-Scan) which contains fewer than 6 edges (either side of three individual peaks). This removes regions to the left of the OoC and BM

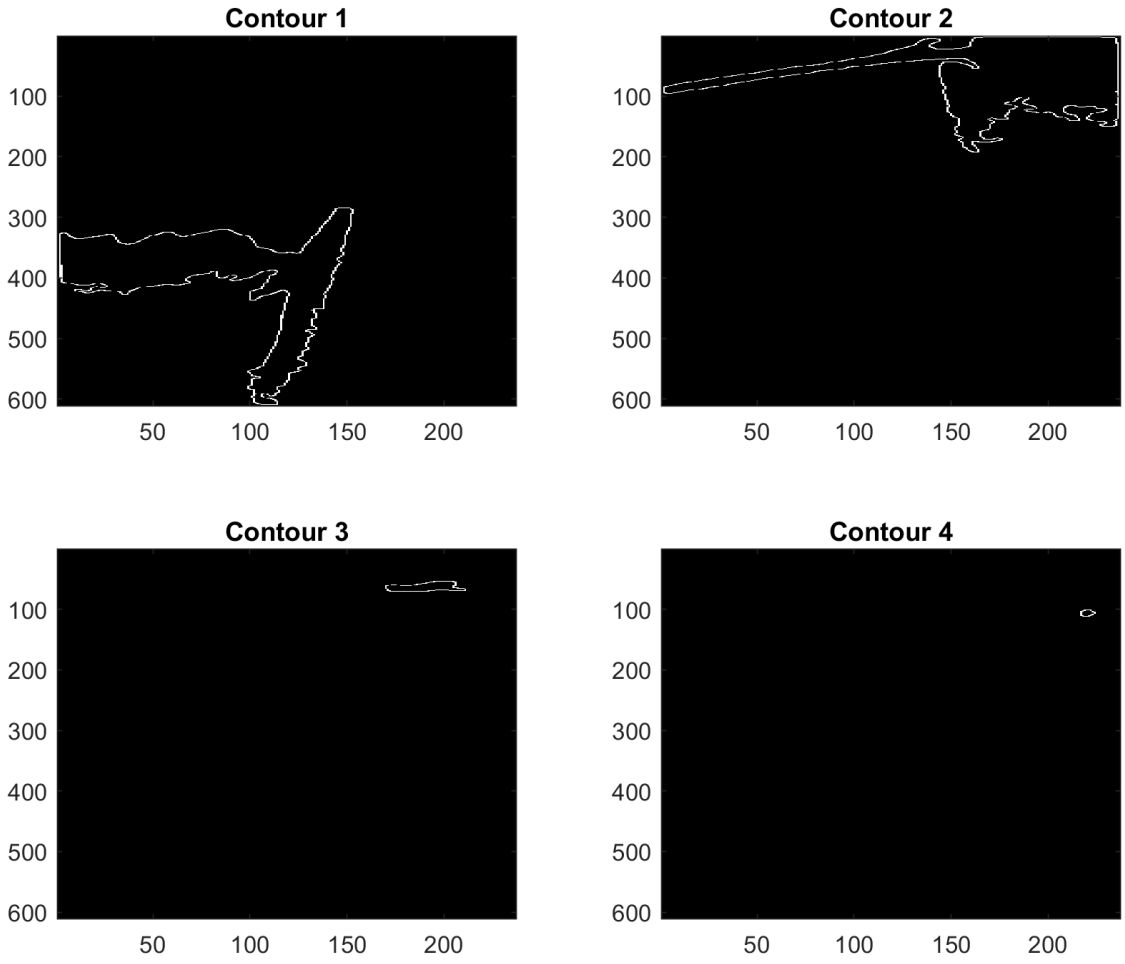

Fig. 7: The four largest closed contours in the edge image. The largest closed contour is consistently that which contains the BM and OHC region, with the round window region coming in second. The other closed contours are mostly uninteresting variations in the bone.

entirely, and some of the bony region. The output of this step can be seen in Figure 8.

H. Tertiary Segmentation

Unfortunately, complex structures within the bone can allow some bone to survive the secondary segmentation. As such, we need one final step to isolate the region of interest. To do so, we realize that the structures within bone lead to significantly smaller gaps than the large gap in the region of interest. As such, we count the size of gaps along each vertical, and remove those vertical lines not containing gaps within the expected (from anatomy) size of the fluid space in the OoC. To do this, we look at the filled contour and use the function "find". This returns all of the indices within either tissue or bone. When the "diff" function is applied to the array of indices, each value larger than 1 constitutes a gap of that value in pixels. On the chance that some random bone fragments survive even this cull, we then look at the largest surviving contiguous stretch of columns. The function "bwareafilt" applied to the surviving column indices returns the largest continuous stretch, and we delete everything else. This leaves us with only the region of interest, as seen in Figure 9.

I. Finding the Best A-Scan

All A-Scans within the segmented region will suffice from a practical perspective, but we want one with high SNR. The noise level across these relatively small regions is more-or-less the same, so we are most interested in the average signal level at each potential A-Scan. We add up the values of the intensities in the original image along each A-Scan only in the segmented region, then divide by the number of segmented pixels. The largest of these averages has the highest signal-to-noise ratio in the region of interest. Across B-Scans, we simply compare these maxima to determine the "best slice" at which

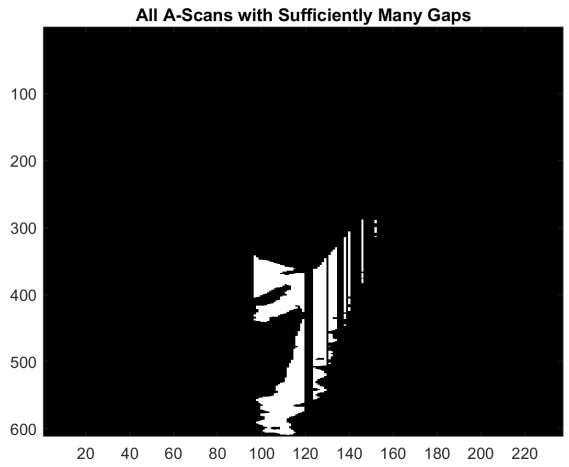

Fig. 8: The output of the second segmentation step, which gives only vertical lines with sufficiently many gaps. The entire irrelevant left-hand side of the first segmentation step output is removed, but some of the bone to the right remains. This must be removed in the third segmentation step.

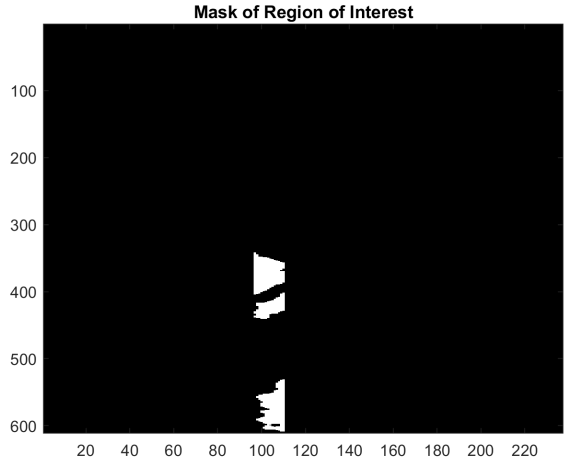

Fig. 9: The mask of the region of interest, containing what is presumably the BM, followed by a gap, then a cellular region in the OoC, and then a larger gap before some bone. This region contains all A-Scans in the original image which could be used for vibration measurements.

to take the A-Scan. Four cross-sections and their determined "best A-Scans" can be seen in Figure 10.

III. Discussion

Unfortunately, due to the COVID-19 pandemic, only one dataset was available for the testing of this software. This was a single volumetric scan from an excised gerbil cochlea, which presents a slightly different anatomy than what one might see *in vivo*. However, there is reason to believe each of these steps would still show success in other samples. The median and Guassian filter steps, for example, we have used on many images with success. The thresholding and edge detection steps are applicable to general images, let alone general images of the cochlea. The segmentation methods are designed only to look for a gap, which is apparent in all cochlea preparations. No deeper anatomical assumptions are ever made in the design of the algorithm.

For this sample, we fnd that we can successfully segment the region of interest in every B-Scan where it the gap is apparent. The maximum SNR A-Scans are, qualitatively, quite good for an excised cochlea, where all reflectivities are lower than they are *in vivo*. Samples from guinea pigs, and *in vivo* samples will need to be explored when they are available, as they may expose faults or lapses in generality that are unclear from this sample.

In terms of time, this algorithm takes about 3 seconds to run on the volumetric scan tested. This volume scan covers the entire round window, meaning that the scan could not get much larger, and this is an upper bound on runtime. The acquisition and processing of a volume scan alone takes about 5 to 6 seconds using the ThorImage software, meaning that the total time incurred by running this algorithm is about 10 seconds at maximum. For comparison, it can take anywhere between 5 minutes and 15 minutes to find a good A-Scan location manually, and this sometimes occurs multiple times during a single experiment. As these experiments are time-sensitive, this would be a very useful tool.

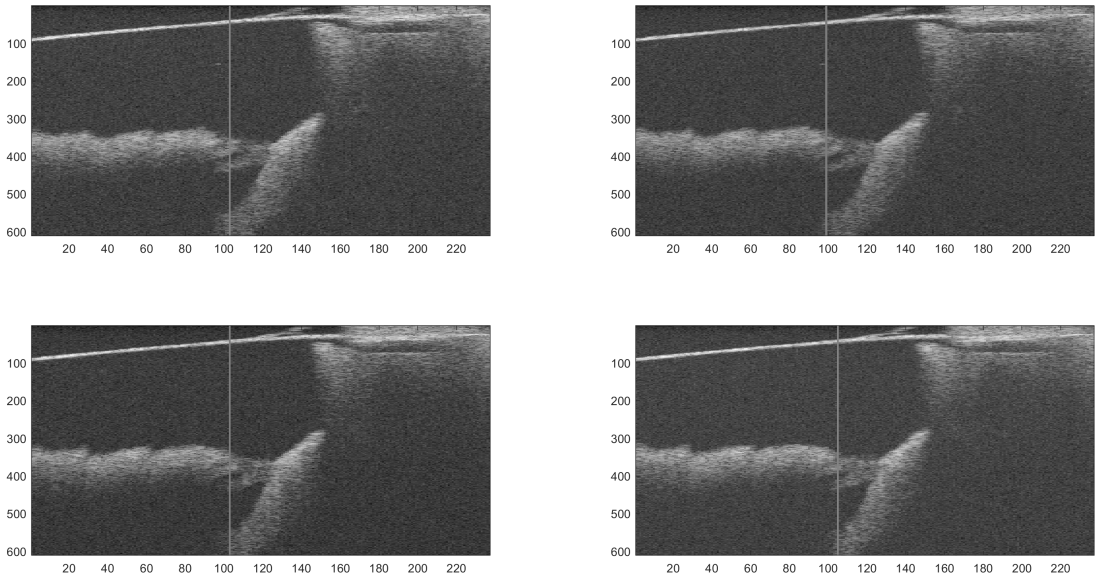

Fig. 10: The result of the algorithm for four consecutive cross-sections, where the gray vertical line in each image represents the location of the optimal A-Scan. Anatomically, it is clear that these all cover the gap that we are interested in.