ORIGINAL ARTICLE

Good and bad boundaries in ultrasound compounding: preserving anatomic boundaries while suppressing artifacts

Alex Ling Yu Hung[1] · **John Galeotti**[2]

Abstract

Purpose Ultrasound compounding is to combine sonographic information captured from different angles and produce a single image. It is important for multi-view reconstruction, but as of yet there is no consensus on best practices for compounding. Current popular methods inevitably suppress or altogether leave out bright or dark regions that are useful and potentially introduce new artifacts. In this work, we establish a new algorithm to compound the overlapping pixels from different viewpoints in ultrasound.

Methods Inspired by image fusion algorithms and ultrasound confidence, we uniquely leverage Laplacian and Gaussian pyramids to preserve the maximum boundary contrast without overemphasizing noise, speckles, and other artifacts in the compounded image, while taking the direction of the ultrasound probe into account. Besides, we designed an algorithm that detects the useful boundaries in ultrasound images to further improve the boundary contrast.

Results We evaluate our algorithm by comparing it with previous algorithms both qualitatively and quantitatively, and we show that our approach not only preserves both light and dark details, but also somewhat suppresses noise and artifacts, rather than amplifying them. We also show that our algorithm can improve the performance of downstream tasks like segmentation.

Conclusion Our proposed method that is based on confidence, contrast, and both Gaussian and Laplacian pyramids appears to be better at preserving contrast at anatomic boundaries while suppressing artifacts than any of the other approaches we tested. This algorithm may have future utility with downstream tasks such as 3D ultrasound volume reconstruction and segmentation.

Keywords Ultrasound image compounding · Ultrasound anatomic boundaries · Ultrasound reconstruction · Laplacian pyramid

Introduction

Even though ultrasound sonography is a low-cost, safe, and fast imaging technique that has been widely used around the globe in clinical diagnosis, surgical monitoring, medical robots, etc., there are still some major drawbacks in ultrasound imaging. Due to the nature of how ultrasound images are captured, it can be hard to see the structures that are deep or underneath some highly reflective surfaces [15]. Certain tissues or structures would bounce back or absorb the sound waves, resulting in dark regions underneath. Such tissues and structures can sometimes produce alterations in ultrasound images which do not represent the actual contents, i.e., artifacts [17]. Moreover, the directionality of ultrasound imaging can make some (parts of) structures difficult to image from certain directions, which may prevent ultrasound images from conveying a complete description of what is going on inside the patient's body. In addition, the directionality may also create confusion for clinicians or medical robots performing downstream tasks. For example, a bullet inside a patient's body would create significant reverberation artifacts that occlude what is underneath. Additionally, when a medical robot inserts a needle into a patient, the reverberation artifacts created by the needle might make the needle tracking algorithm fail or disrupt the identification of the structures of interest [24]. Even though some artifacts have diagnostic significance, which could help clinicians localize certain

✉ Alex Ling Yu Hung
lingyuh@andrew.cmu.edu

John Galeotti
jgaleotti@cmu.edu

1 Biomedical Engineering Department, Carnegie Mellon University, Pittsburgh, PA 15213, USA

2 Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA

structures or lesions inside patients' bodies [1,2], the artifacts become less meaningful once the objects of interest are identified. Furthermore, if we preserve the artifacts from different viewpoints, then they could substantially occlude real tissues and the image would be harder to interpret. When there are multiple viewpoints available in ultrasound imaging, we can reconstruct an ultrasound image that represents the underlying structures better while having fewer artifacts.

However, no existing method can do the job perfectly. Relatively simple methods such as averaging the overlapping pixel values from different viewpoints [33] or taking the maximum of such pixels [18] result in lower dynamic range or additional artifacts in the output images. Other more advanced ultrasound compounding algorithms [7,10] reconstruct the 3D volume of ultrasound using a tensor representation, but both of them still combine the overlapping pixels by simply averaging them or taking the maximum. The method proposed by zu Berge et al. [4] utilizes the per-pixel confidence map proposed by Karamalis et al. [16] as weights in compounding. While this method does not directly take the average or maximum, it does not take the contrast of the image into account. In the survey paper by Mozaffari et al. [21], a large number of 3D compounding methods are covered, but all of the methods deal with the overlapping pixels from different views by taking the average or maximum. Since both bright and dark regions contain useful information in ultrasound images, maximizing the contrast in those regions while lowering the intensity of noise and artifacts in other regions is essential in compounding. In all cases, every existing compounding algorithm tends to introduce new artifacts into the volumes and lower the dynamic range. These algorithms can only preserve either dark or bright regions, but in some clinical settings or in computer vision algorithms to guide downstream medical robots, *both dark and bright regions are useful*.

We do not want to naively take the maximum or average when dealing with overlapping pixels from different views, since doing so would lower the contrast or create new artifacts. Our goal is to clearly differentiate all structures, whether dark or bright, while suppressing artifacts and speckle noise to help with downstream computer vision tasks such as vessel segmentation. To better reconstruct the vessels and detect the bones, unlike prior work, we are less concerned with recovering the most "accurate" individual pixel values but more concerned with enhancing the images by maximizing the contrast. We focus on preserving patches with the largest contrast, suppressing less-certain high-frequency information to prevent piecemeal-stitching artifacts and reduce existing artifacts. Our most important contributions are: (1) Use more advanced methods when compounding overlapping pixels between different views instead of directly taking the average or maximum. (2) Keep the pixels and structures with the higher confidence when

compounding. (3) Preserve the pixels or patches that have the largest local contrast among the overlapping values from different viewpoints. (4) Identify anatomic boundaries of structures and tissues and treat them differently while compounding. (5) Use Laplacian pyramid blending [5] to remove discrepancy in pixel values from ultrasound images captured in different viewpoints. (6) Make use of the advantages of different compounding methods in different frequency scales.

Related work

As for freehand ultrasound compounding, in 1997, Rohling et al. [26] proposed to compound the freehand ultrasound images in the same plane iteratively, using an approach that is based on averaging. Later, the same group used interpolation to reconstruct 3D volumes of non-coplanar freehand ultrasound and still averaged the overlapping pixels [27]. As mentioned by Rohling et al. [28] and Mozaffari et al. [21], the most common method in freehand compounding is to use interpolation to calculate the missing pixels and use averaging to calculate the overlapping pixels, while this might not be the best approach. Grau et al. [9] came up with a compounding method based on phase information in 2005. Although this method is useful, access to radio frequency (RF) data is limited, preventing the algorithms from being widely adopted. Around the same time, Behar et al. [3] showed that averaging the different view worked well if the transducer were set up in a certain way by simulation, but in practice, it would be extremely hard to set up the imaging settings that way.

In recent years, Karamalis et al. [16] came forward with a way to calculate physics-inspired confidence values for each ultrasound pixel using a graph representation and random walk [8], which Zu Berge et al. [4] used as weights in a weighted average algorithm to compound ultrasound images from different viewpoints. Afterward, Hung et al. [14] proposed a new way to measure the per-pixel confidence based on directed acyclic graphs that can improve the compounding results. Hennersperger et al. [10] and Göbl et al. [7] modeled the 3D reconstruction of ultrasound images based on more complete tensor representations, where they modeled the ultrasound imaging as a sound field. While these two recent papers made great advances in reconstructing ultrasound 3D volumes, they still compound overlapping pixels by averaging or taking the maximum. A review of freehand ultrasound compounding by Mozaffari et al. [21] summarized compounding methods using 2D and 3D transducers. However, few papers talked about how they deal with overlapping pixels, which is what our work mainly focuses on.

In the case of robot control instead of freehand ultrasound, Virga et al. [34] modeled the interpolation/inpainting problem as partial differential equations and solved them

with a graph-based method purposed by Hennersperger et al. [11,34]. They also did image compounding based on the tensor method by Hennersperger et al. [10].

Although ultrasound artifacts have barely been directly considered in previous compounding approaches, it has been widely discussed in literature. Reverberation artifacts and shadowing are useful in diagnosis because those artifacts can help clinicians identify highly reflective surfaces and tissues with attenuation coefficients significantly different from normal tissues [12]. Reverberation artifacts are most useful in identifying anomalies in lungs [2,29], while it can also be used in thyroid imaging [1]. Shadowing could be used in measuring the width of kidneys [6]. However, artifacts and noise could occlude the view of other objects of interests [20] or hurt the performance of other tasks, such as registration [25], needle tracking [24], or segmentation [36]. In recent years, several learning-based methods have been focusing on identifying artifacts and shadows and using this information to identify other objects [13,19], but they all need substantial labeling work and a relatively large dataset. Non-learning-based methods to remove the artifacts either use RF data [31] or temporal data [35] or fill in the artifact regions based on neighboring image content within the same image [30]. All of these methods make assumptions about what the missing data probably look like, whereas our approach utilizes multi-view compounding to obtain actual replacement data for the artifact pixel locations.

Methods

Identifying good boundaries

Any sort of averaging between different views in which an object appears either too bright or dark in one view will lower the compounded object's contrast with respect to surrounding pixels. Even though artifacts could be suppressed, the useful structures would also be less differentiated, which is not the optimal approach. Therefore, identifying good anatomic boundaries, and treating them differently than other pixels in compounding, is essential to preserving the dynamic range and contrast of the image.

Ultrasound transmits sound waves in the axial (e.g., vertical) direction, so sound waves are more likely to be bounced back by horizontal surfaces. Horizontal edges are also more likely to be artifacts, in particular reverberation artifacts [23]. The trait of reverberation artifacts is that the true object is at the top with the brightest appearance compared to the lines beneath which are artificial. The distance between the detected edges of reverberation artifacts is usually shorter than other structures. Also, structures in ultrasound images are usually not a single line of pixels: They usually have thickness. Though reverberation artifact segmentation algorithms

like [13] could work well in identifying the bad boundaries, labeling images is a very time-consuming task. Besides, the exact contour of the structures in ultrasound images is ambiguous, which can be hard and time-consuming to label as well, so directly using manual labels would be less efficient and it might introduce new artifacts into the images. Therefore, we propose to refine the detected edges based on the appearances of reverberation artifacts.

First, we detect the horizontal boundaries through edge detection algorithms. To detect the actual structures in the ultrasound images instead of the edge of the structure, we calculate the gradient at pixel (x, y) by taking the maximum difference between the current pixel and α pixels beneath,

$$\frac{\partial I(x, y)}{\partial y} = \max_{j=1,2,\ldots,\alpha} |I(x, y) - I(x, y + j)| \tag{1}$$

where in this paper, we set α to 15.

We then group the pixels that are connected into clusters, such that pixels belonging to the same boundary are in the same cluster. We remove the clusters containing fewer than 50 pixels. After that, we only keep the clusters that do not have a cluster of pixels above itself in β pixels. In this paper, $\beta = 20$.

A refinement is performed by iterating through the kept clusters and comparing the pixel values against that of the original image. A stack s is maintained, and the pixels in the kept clusters with values greater than $threshold1$ are pushed into it. We pop the pixel (x, y) at the top of the stack and examine the pixels in its 8-neighborhood (x_n, y_n). If (x_n, y_n) has never been examined before and satisfies $I(x_n, y_n) > threshold1$ and at the same time the gradient value is less than $threshold2$, i.e., $|I(x_n, y_n) - I(x, y)| < threshold2$, then we push (x_n, y_n) into the stack s. We repeat this procedure until s is empty. We add this step because the boundary detection might not be accurate enough, and we can ignore detected boundaries with low pixel values to suppress false positives. In this paper, $threshold1$ and $threshold2$ are set to 30 and 2, respectively. The pseudocode for the described algorithm is shown in Algorithm 1. We note that we assigned the values to the parameters based on empirical results.

Compounding algorithm

Attenuation reduces ultrasound image contrast in deeper regions. Simply taking the maximum, median, or mean while compounding [18] further undermines the contrast information, where structure information is stored. Taking the maximum also would create artifacts by emphasizing non-existent structures resulting from speckle noise in uncertain regions. Although uncertainty-based compounding approach by [4] suppresses the artifacts and noise to some extent, it produces substantially darker images than the originals and

---

Algorithm 1: Horizontal-edge refinement

Data: input image I
Result: output boundary mask B
$edges = clustering(denoising(\frac{dI}{dy}))$;
$edges = cleanup(edges)$;
$mask = zeros(I.shape)$; $B = zeros(I.shape)$;
for $edge$ in $edges$ **do**
   if $\forall e$ in $edges$ that are far enough from edges underneath
   then
      | $mask[edge] = 1$;

for $[i, j]$ where $mask[i, j] == 1$ **do**
   if $B[i, j] == 0$ **then**
      stack s; # initialize stack
      if $I[i, j] > threshold1$ **then**
         | $s.push([i, j])$; $B[i, j] = 1$;
      while s is not empty **do**
         $[x, y] = s.pop$;
         for $[ii, jj]$ in the neighborhood of $[x, y]$ **do**
            if $I[ii, jj] > threshold1$ and
            | $I[x, y] - I[ii, jj] | < threshold2$ and
            $B[ii, jj] == 0$ **then**
              | $s.push([ii, jj])$; $B[ii, jj] = 1$;

---

lowers the dynamic ranges. Also, taking the maximum retains the bright regions, but some dark regions are also meaningful, so it would make more sense to preserve the patches with the largest local contrast than to simply select the pixels with maximum values. However, directly taking pixels with the largest contrast would lead to neighboring pixels inconsistently alternating between different source images. Besides, the neighbors of a pixel might all be noise, resulting in instability of the algorithm. Taking the maximum contrast might also emphasize the artifacts.

We developed a novel Laplacian pyramid [5] approach to compound the images at different frequency bands and different scales. In this way, we can apply contrast maximization method at certain frequency bands while reconstructing from the pyramid. However, the pixels at extremely large scale in the pyramid represent a patch containing a huge number of pixels in the lower layers, so the contrast in this layer has less anatomical meaning. On the other hand, when the scale is small, the noise in the image would create large local contrast, so maximum weighted contrast might introduce new artifacts into the image. At extremely low and high scales, we thus consider contrast to be less important than intensity confidence measures. Another flaw of directly maximizing the contrast is that the large contrast region might contain artifacts and shadows, so we only maximize the contrast when the overlapping pixels have similar structural confidence values [14]; otherwise, we use the pixel with the larger structural confidence value in the compounded image, as low structural confidence value indicates that the pixel belongs to artifacts or shadows. Although some anatomic structures would be removed due to the low confidence values, artifacts and noises would also be removed in the compounded image. The anatomic structures are later compensated for in the later stage of the algorithm.

As is shown in Fig. 1, our novel ultrasound compounding method takes ultrasound images from multiple viewpoints and calculates their intensity and structural confidence maps [14] and then calculates Laplacian and Gaussian [32] pyramids of the original images and the Gaussian pyramid of confidence maps. Denote $L_{m,n}$ $GI_{m,n}$ as the nth layer of the Laplacian pyramid and Gaussian pyramid of the mth coplanar ultrasound image, respectively, $GC_{m,n}$ $G\Gamma_{m,n}$ as the nth layer of the Gaussian pyramid of the intensity and structural confidence map of mth coplanar ultrasound image, respectively, and L_k as the kth layer of the Laplacian pyramid of the synthetic image. M is the set of viewpoints, with $|M|$ views. Also, denote $N(i, j)$ the 8-connected neighborhood of pixel (i, j). Here, we combine the weighted maximum contrast and weighted average together. For the kth layer of the pyramid, if the difference across viewpoints between the maximum and minimum structural confidence values $G\Gamma_{m,k}(i, j)$, where $m \in M$, is less than a certain threshold γ ($\gamma = 0.05$ in this paper), we take the pixel (i, j) with the largest contrast at this scale, since only when there is no artifact at the pixel, does taking the largest contrast make sense

$$\widetilde{m}(i, j) = \arg\max_{m \in M} \sum_{(a,b) \in N(i,j)} | GI_{m,k}(a, b) - GI_{m,k}(i, j) | \tag{2}$$

If not, we take the pixel (i, j) with the largest structural confidence at this scale

$$\widetilde{m}(i, j) = \arg\max_{m \in M} G\Gamma_{m,k}(i, j) \tag{3}$$

Denote the intensity-confidence-weighted average at the kth layer of the Laplacian pyramid as La_k,

$$La_k(i, j) = \frac{\sum_{m=1}^{|M|} GC_{m,k}(i, j) L_{m,k}(i, j)}{\sum_{m=1}^{|M|} GC_{m,k}(i, j)} \tag{4}$$

Then, the kth layer of the Laplacian pyramid of the synthetic image can be calculated as,

$$L_k(i, j) = \phi(k) L_{\widetilde{m}(i,j),k}(i, j) + (1 - \phi(k)) La_k(i, j) \tag{5}$$

where

$$\phi(k) = \frac{1}{0.4\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{(2k-K-1)^2}{0.16(K-1)^2}\right)} \tag{6}$$

is a weight function, and K is the number of total layers. This weight function is designed to assign lower weights to contrast maximization and higher weights to intensity-confidence-weighted average in extremely low and high scale.

The compounding algorithm could be further generalized:

$$L_k(i, j) = \sum_{n=1}^{N} \phi_n(k) F_n(\{L_{m,k}\}_{m \leq |M|}, \{G_{m,k}\}_{m \leq |M|}) \quad (7)$$

where

$$\sum_{n=1}^{N} \phi_n(k) = 1, 0 < k \leq K, 0 < n \leq N, 0 \leq \phi_n(k) \leq 1 \quad (8)$$

K is the total number of layers, N is the total number of compounding methods, p is the total number of viewpoints, $G_{m,k}$ denote any kind of confidence map at layer k from viewpoint m, and F_n denote a compounding method.

We can use any weighting scheme to combine any number of compounding schemes in the Laplacian pyramid based on the application and data.

The algorithm still takes some sort of confidence-based weighted averaging in some layers of the pyramid. During artifact-free contrast maximization, some anatomic boundaries would be removed incorrectly due to lower structural confidence. Therefore, even though this approach works well in preserving contrast and suppressing artifacts, the actual boundaries of structures still tend to get darker. In addition to what we just proposed above, the algorithm we purposed back in section "Identifying good boundaries" can also be incorporated. While reconstructing the image from the new Laplacian pyramid after getting the image from the third layer, the good boundaries are detected and values from the original images are taken. For overlapping pixels here, we take the maximum. We apply the same notation as above, and $GB_{m,k}$ is layer k from viewpoint m of the Gaussian pyramid of the boundaries mask B (Gaussian pyramid of algorithm 1's output).

$$L_3(i, j) = \max \left(\frac{\sum_{m=1}^{|M|} GB_{m,3}(i, j) GI_{m,3}(i, j)}{\sum_{m=1}^{|M|} GB_{m,3}(i, j)}, L_3(i, j) \right) \quad (9)$$

This step is done on the third layer of the pyramid since there are still two layers before the final output, so piecemeal-stitching artifacts can still be suppressed. The step is not done in deeper layers, so that we can still preserve contrast. The pipeline for combining two individual compounding methods and boundaries enhancement is shown in Fig. 2.

Experiments

Data acquisition

The data used in these experiments were gathered from three different sources: a Advanced Medical Technologies anthropomorphic Blue Phantom (blue-gel phantom), an ex vivo lamb heart, as well as a live pig.

For our initial blue-gel phantom experiments, a UF-760AG Fukuda Denshi diagnostic ultrasound imaging equipment with a linear transducer (51 mm scanning width) set to 12 MHz, a scanning depth of 3 cm, and a gain of 21 db is used to scan the surface of the phantom. A needle is rigidly inserted and embedded within the phantom. When scanning the surface, images from two orthogonal viewpoints are collected. As the phantom square, it is easy to ensure coplanar orthogonal views with using freehand imaging without any tracking equipment. The experiment setup is shown in Fig. 3a.

For the ex vivo experiment, a lamb heart is placed within a water bath in order to insure good acoustic coupling. Using a Diasus High Frequency Ultrasound machine, a 10-22 MHz transducer is rigidly mounted onto a 6 degrees of freedom (dof) Universal Robotics UR3e arm. Using a rough calibration to the ultrasound tip, the 6-dof arm is able to ensure coplanar views of the ex vivo lamb's heart. The experiment setup is shown in Fig. 3b.

For the in vivo experiment, a live pig is used as the imaging subject. A UF-760AG Fukuda Denshi diagnostic ultrasound imaging equipment with a linear transducer (51 mm scanning width) set to 12 MHz, a scanning depth of 5 cm, and a gain of 21 db is mounted on the end-effector of the UR3e arm and is placed on the desired location manually to get a good view of the vessel [37]. Some manual alignments are needed for the arm to be in proper contact with the pig's skin. This pose of the robot is the zero degree view of the vessel. After this, the rotational controller rotates the probe along the probe's tip by the specified angle. For this experiment, we cover a range from 20 degree to -25 degree at an interval of 5 deg. The input to the UR3e robot is sent through a custom GUI that is designed to help the users during surgery. The GUI has relevant buttons for the finer control of the robot in the end-effector frame. The GUI also has a window that displays the ultrasound image in real time which helps in guiding the ultrasound probe.

Qualitative evaluation

We visually compare the results of our method against average [33], maximum [18], and uncertainty-based fusion [4]. As is shown in Fig. 4, our algorithm has the best result in suppressing artifacts, and at the same time, the brightness of the boundaries (green arrows) from our algorithm is similar to that of taking maximum [18]. Our method also preserves
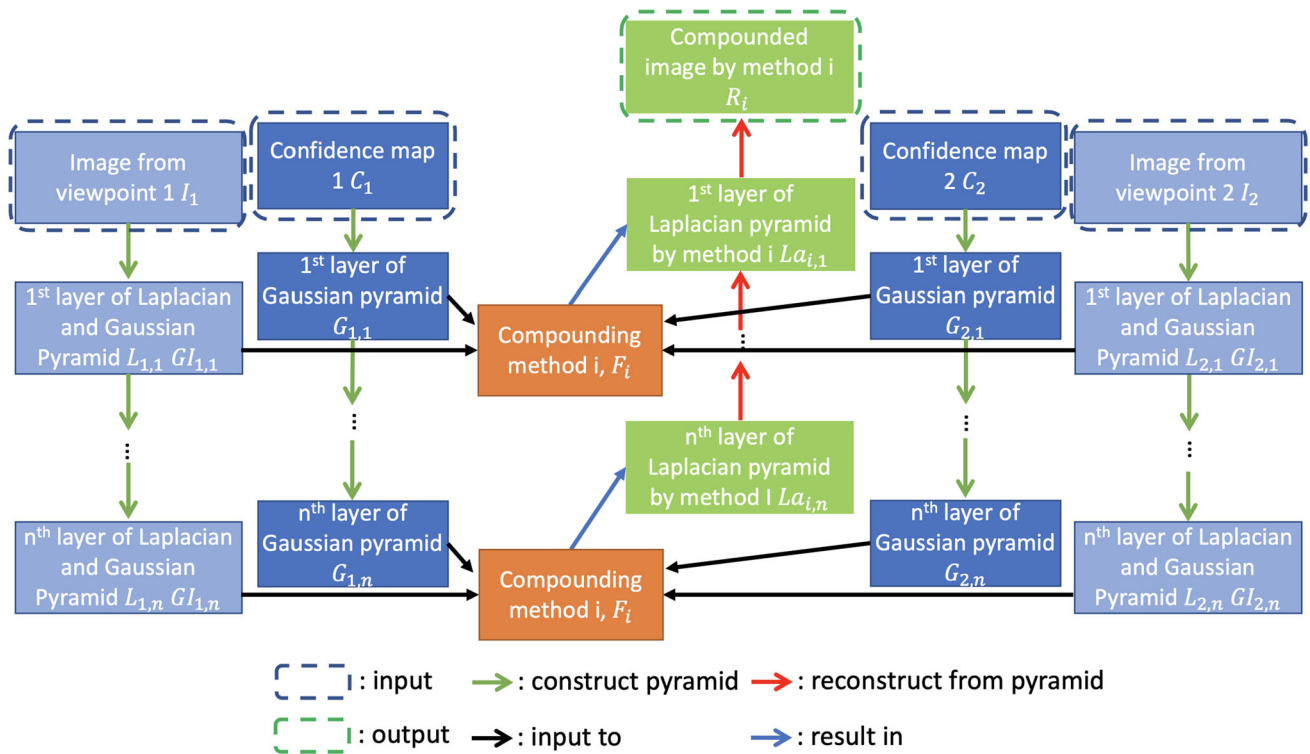
Fig. 1 Compounding with Laplacian and Gaussian pyramid. The compounding is performed in each layer of the pyramid with the confidence map (intensity confidence or structural confidence) used as some sort of weights. The compounding results are reconstructed from the pyramid of the compounded image
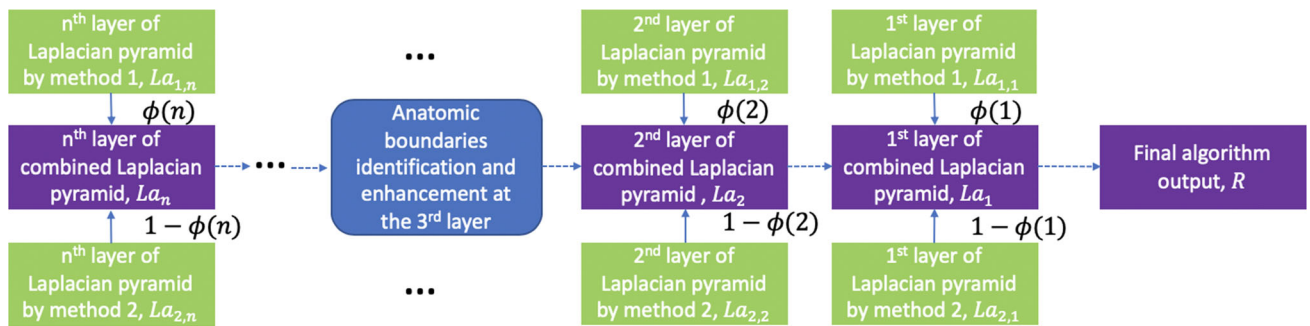
Fig. 2 Pipeline for combining two individual compounding methods and boundaries enhancement. We combine the results from different methods by different weights in each layer of the pyramid. The anatomic boundaries are enhanced at the third layer so that the enhancement does not introduce new artifacts

a lot more contrast since other parts of the patch are darker in comparison with our bright boundaries, whereas the boundaries from the other two compounding algorithms are darker and therefore less contrasting with the dark interior. Our algorithm also completely suppresses the reverberation artifacts in the regions that the red and yellow arrows point to, while the results from other algorithms all preserve undesirable aspects of artifacts.

To compare our results against other existing compounding algorithms (average [33], maximum [18], and uncertainty-based fusion [4]), we select 5 examples of results on the anthropomorphic phantom, which is shown in Fig. 5.

In the first row, our algorithm almost completely removes the reverberation artifacts in the synthesized image and at the same time preserves the contrast in the images. In other phantom examples, our algorithm is also the best in removing the reverberation artifacts and shadows the vessel walls cast, while preserving the brightness of the vessel walls, needles, and other structures in the images. Our algorithm preserves the "good boundaries" that represent the anatomic boundaries while suppressing boundaries that are not real.

Besides, we also test our algorithm on real tissue images. The comparison between our algorithm and other existing algorithms on the lamb heart is shown in Fig. 6, where only
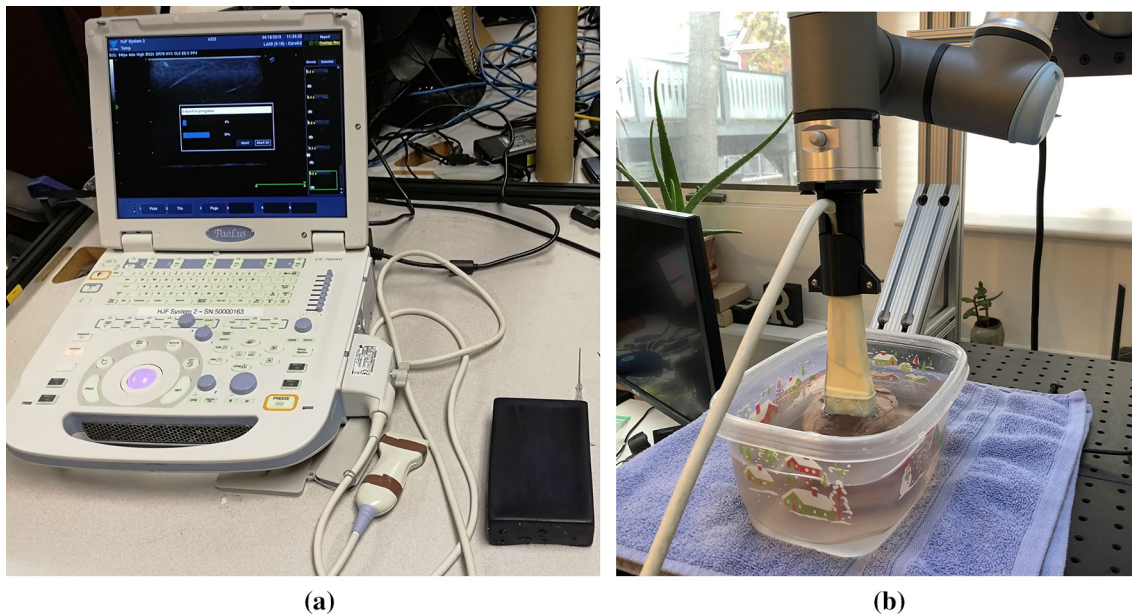
(a) **(b)**

Fig. 3 **a** The experiment setup for the blue-gel phantom experiment. The square phantom and the needle are shown in the image. We perform the experiment with freehand imaging since it is easy to make sure the orthogonal views on a square phantom. **b** The experiment setup for the lamb heart experiment, where the lamb heart is situated in a water bath to ensure acoustic coupling. The imaging is done by a robot-controlled high-frequency probe
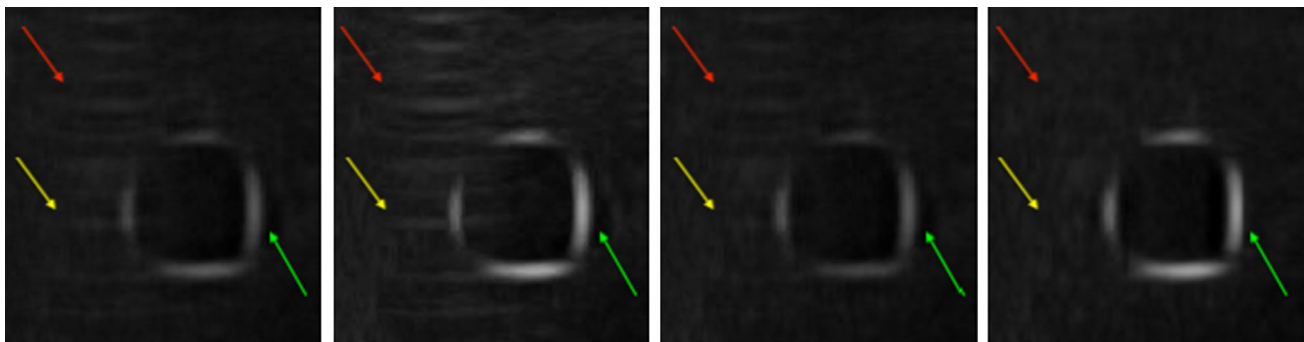

Fig. 4 Compounded patches left to right: average [33], maximum [18], uncertainty-based fusion [4], and our algorithm, where the green arrows indicate the vessel walls, while the red and yellow arrows indicate the artifacts. As is shown in the figure that our result preserves the brightness of the vessel boundaries and suppresses the artifacts at the same time, while other methods fail to do so

maximum [18] and our algorithm preserve the contrast at the red arrows, whereas the results by other algorithms are darker in that patch. However, maximum [18] fails to preserve the contrast at the blue arrows, while our method keeps the contrast at both red and blue arrows. It also shows that even on highly noisy data, our algorithm also has decent performance.

We further demonstrate that our method is able to handle images from more than two viewpoints, i.e., situation where $|M| > 2$. In this experiment, we utilize the live-pig data and instead of using the structural confidence, we use a simple contrast maximization (equivalent to the case when all structural confidence at corresponding pixels is equal), due to the difficulty in getting the reference image for the structural con-

fidence. The result is shown in Fig. 7, where the change in probe position between the first two images only consists of translation, and when moving the probe to the third location, it also involves rotation.

We also would like to show how each component of our algorithm contributes to the final output. Our proposed algorithm mainly consists of three parts: (1) structural confidence-based artifact-free contrast maximization, (2) intensity confidence-based weighted averaging, and (3) edge enhancement. As shown in Fig. 8, structural confidence-based artifact-free contrast maximization (1) removes the reverberation artifacts and shadows decently well and preserves some contrast in the images, but some parts of the

Fig. 5 Results on the phantom with a needle inserted in it. The left two columns are the two input images. (Phantom images were acquired orthogonally within plane, where the imaging direction of the first and second column is from left to right and from top to bottom, respectively.) The right four columns from left to right are results from average [33], maximum [18], uncertainty-based fusion [4], and our algorithm. On the phantom examples, it is clear that our method best preserves bright and dark anatomy while suppressing artifacts

vessel boundaries are removed as well and create some unnatural holes in the image. Intensity confidence-based weighted averaging (2) preserves the vessel boundaries but not as bright as before, and it also removes the reverberation artifacts but also not as good as structural confidence-based artifact-free contrast maximization (1). Edge enhancement (3) clearly enhances the boundaries but at the same time slightly enhances a small portion of the reverberation artifacts (yellow arrow) as well. Generally, the final output ((1)(2)(3)) leverages the different components of the image, having less reverberation artifacts than the result by using only (2) and (3) (red arrow), while having no irregular holes like the result by only (1) and (3) (blue arrow). Depending on different application, we can adjust the weights $\phi(k)$ and how we utilize

the detected good boundaries, to compound the images in the way we want.

Quantitative evaluation

We continue to compare our results with average (avg) [33], maximum (max) [18], and uncertainty-based fusion (UBF) [4], as well as the original images. The challenges to evaluate the results are: (1) There are no ground truth images that show what the compounded images should look like, (2) our algorithm is designed to maximize the contrast near boundaries and suppress the artifacts, so the exact pixel values do not matter, so manually labeled binary masks where anatomic boundaries are 1 and other pixels are 0 would not
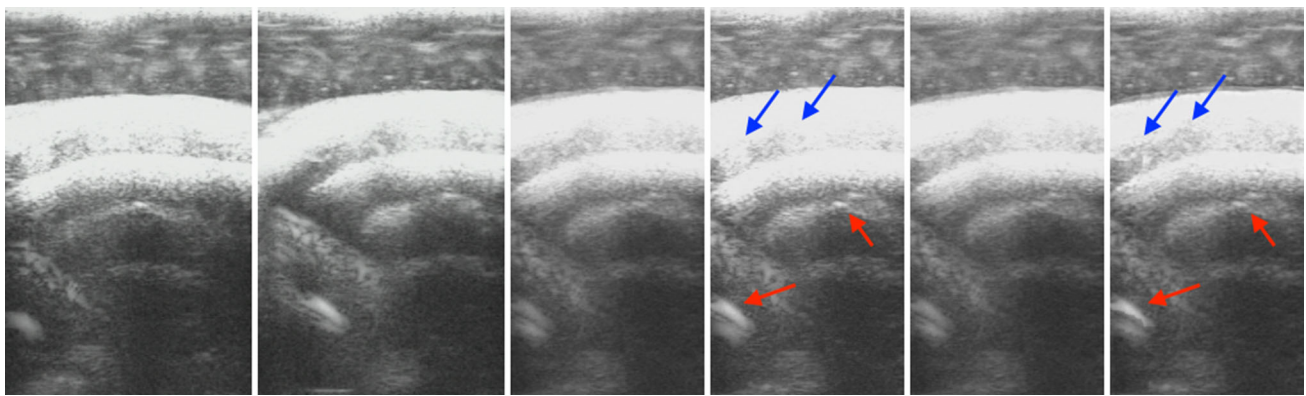
Fig. 6 Results on the lamb heart ultrasound images. From left to right: two input images, compounded results (only overlapped regions are shown) by average [33], maximum [18], uncertainty-based fusion [4], and our algorithm. Maximum [18] and ours the only methods that are able to preserve the bright boundaries at the red arrows, but the maximum is not able to preserve the contrast at the blue arrows like ours does
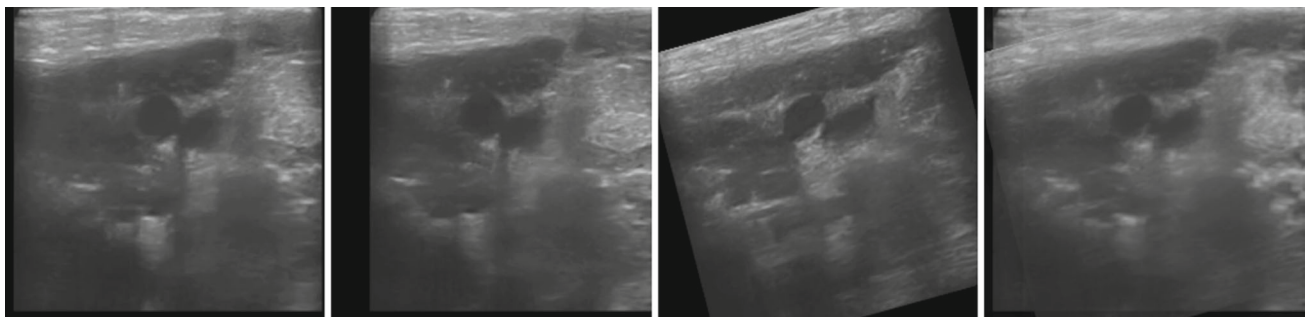

Fig. 7 The compounding result of three live-pig images. The left three images are the input images, while the right image is the result. In the result, the vessel and the structure on the right are successfully preserved, while the shadows cast by the vessel become less significant

work as some naive ground truths. Besides, since the majority of the pixels would be 0 in those naively labeled images, the peak signal-to-noise ratio (PSNR) with such images as ground truth would be a lot larger if the images are dark compared with images with larger visual contrast.

To show that our method generates images with better quality, we propose to use our variance-based metric. We separately evaluate image patches containing artifacts, which should have low contrast, and patches containing boundaries, which should have high contrast. For the patches with artifacts, we evaluate the algorithms based on the ratio between the variance of the patch and the variance of the whole image (denoted as variance ratio), as well as the ratio between the mean of the patch and the mean of the whole image (denoted as mean ratio). The patches with the artifacts should have lower variance and a similar mean compared with the whole image, since artifacts are supposed to be suppressed. As for patches with real boundary signals, we only care about the contrast, so our metric is the variance ratio. We want the variance in the patches with boundary signals to be much larger than the variance of the whole image. We compute the average mean ratio (AMR) and average variance ratio (AVR)

Table 1 Evaluation by mean and variance

	View1	View2	Avg	Max	UBF	Ours
Artifacts						
AMR	1.434	1.996	1.757	1.433	1.277	**1.206**
AVR	0.109	0.224	0.204	0.234	0.134	**0.048**
Boundaries						
AVR	3.609	2.172	3.007	3.612	1.696	**3.876**

Bold numbers indicate best performance

on 27 signal patches and 23 artifact patches. These patches are cropped from the same position in every image to keep the comparison fair, and examples of the patches are shown in Fig. 9, where the green boxes are the anatomic boundary signal patches and the red boxes are the artifact patches. The results are listed in Table 1. Our method outperforms other algorithms in suppressing artifacts. As for real boundary signals, our method appears superior to all the other methods.

Additionally, we also compare our results with the previous ones by performing vessel segmentation on the compounded images. We manually selected 6 patches containing vessels with annotated vessel boundaries as ground truth. We
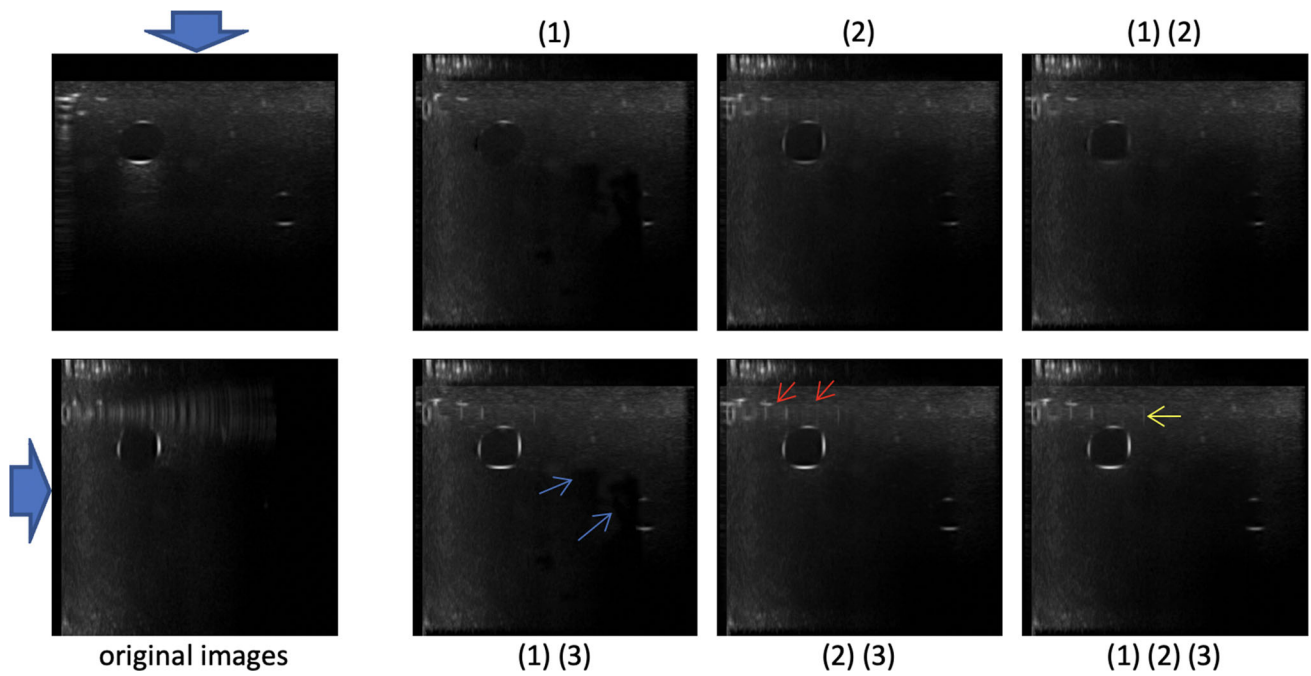
Fig. 8 Results showing the effect of different parts of the algorithm. The left column consists of the original images with arrows indicating the imaging direction. The right 6 images are compounding results where the numbers above or under the images indicate which part(s) of the algorithm is (are) used to constructed the compounded images. Note that the correspondence of the numbers is (1) structural confidence-based artifact-free contrast maximization, (2) intensity confidence-based weighted averaging, (3) edge enhancement
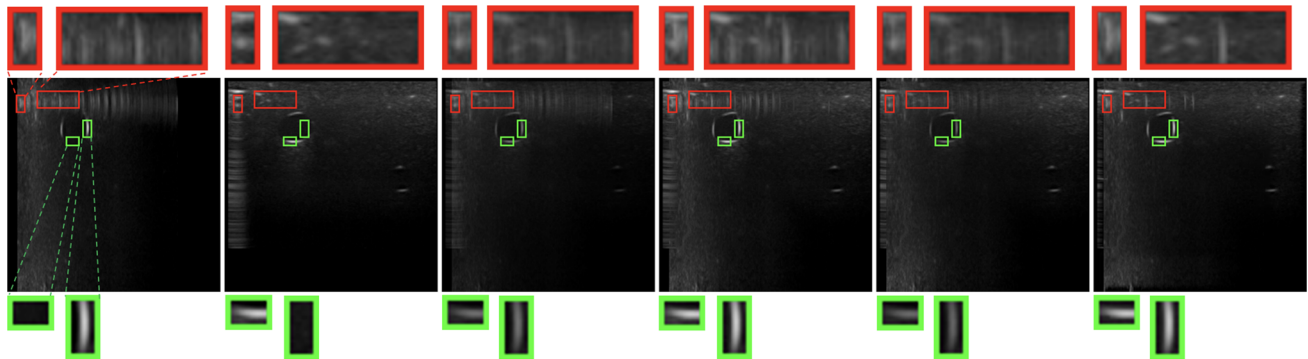
Fig. 9 An example of boundary signal and artifact patches. From left to right are two original images that are orthogonal, followed by results from average [33], maximum [18], uncertainty-based fusion [4], and our algorithm. The green and red boxes are examples of boundary and artifact patches

Table 2 Evaluation by segmentation

	View1	View2	Avg	Max	UBF	Ours
Dice coefficient	0.654	0.575	0.673	0.719	0.594	**0.737**

Bold numbers indicate best performance

perform a naive segmentation as the following. We first apply Otsu thresholding [22] to each compounded patch which contains blood vessels to automatically separate the vessel boundaries from the background. We then fit an ellipse to the separated boundary points by the first step to segment the vessel.

Table 2 shows Dice coefficients [38] comparing each method against ground truth where ours have the best performance. Figure 10 shows an example of the segmentation result. This simple adaptive segmentation performs the best on our compounding results. Since Otsu thresholding is purely pixel intensity-based thresholding without considering other information, it is somewhat sensitive to the intensity of noise in the image. Therefore, better segmentation results show that our method is better than the prior algorithms at preserving the vessel walls while suppressing noise and artifacts.
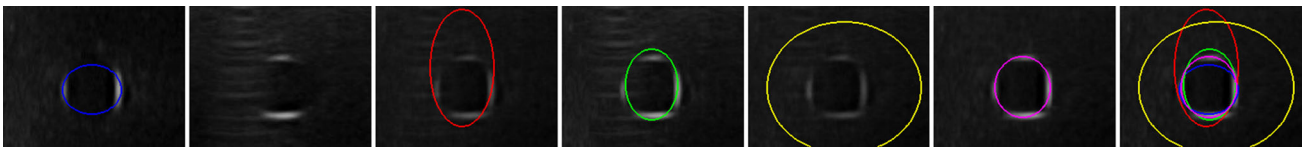
Fig. 10 The result of vessel segmentation. First two images: two original images (the algorithm is not able to fit an ellipse on the second image). Following four images: results by average [33], maximum [18], uncertainty-based fusion [4], and our algorithm. The last image: segmentation results overlaying on the compounded image synthesized by our algorithm. It can be seen that the segmentation on the first original image is flatter because of the missing top and down boundaries, while the segmentation on the result by maximum is affected by the reverberation artifact at the top. Other segmentation results are clearly off, while the segmentation algorithm fits the vessel boundaries very well on our result

Conclusion

In this work, we present a new ultrasound compounding method based on ultrasound per-pixel confidence, contrast, and both Gaussian and Laplacian pyramids, taking into account the direction of ultrasound propagation. Our approach appears better at preserving contrast at anatomic boundaries while suppressing artifacts than any of the other compounding approaches we tested. Our method is especially useful in compounding problems where the images are severely corrupted by noise or artifacts, and there is substantial information contained in the dark regions in the images. We hope our method could become a benchmark for ultrasound compounding and inspire others to build upon our work. Potential future work includes 3D volume reconstruction, needle tracking and segmentation, artifact identification and removal, etc.

Declarations

Conflict of interest Galeotti serves on the advisory board for Activ Surgical, Inc., and he is a Founder and Director for Elio AI, Inc.

Ethical approval All applicable international, national, and/or institutional guidelines for the care and use of animals were followed.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Ahuja A, Chick W, King W, Metreweli C (1996) Clinical significance of the comet-tail artifact in thyroid ultrasound. J Clin Ultrasound 24(3):129–133
2. Baad M, Lu ZF, Reiser I, Paushter D (2017) Clinical significance of us artifacts. Radiographics 37(5):1408–1423
3. Behar V, Nikolov M (2006) Statistical analysis of image quality in multi-angle compound imaging. In: IEEE John Vincent Atanasoff 2006 international symposium on modern computing (JVA'06). IEEE, pp 197–201
4. zu Berge CS, Kapoor A, Navab N (2014) Orientation-driven ultrasound compounding using uncertainty information. In: International conference on information processing in computer-assisted interventions. Springer, pp 236–245
5. Burt PJ, Adelson EH (1983) A multiresolution spline with application to image mosaics. ACM Trans Graph (TOG) 2(4):217–236
6. Dunmire B, Harper JD, Cunitz BW, Lee FC, Hsi R, Liu Z, Bailey MR, Sorensen MD (2016) Use of the acoustic shadow width to determine kidney stone size with ultrasound. J Urol 195(1):171–177
7. Göbl R, Mateus D, Hennersperger C, Baust M, Navab N (2018) Redefining ultrasound compounding: Computational sonography. arXiv preprint arXiv:1811.01534
8. Grady L (2006) Random walks for image segmentation. IEEE Trans Pattern Anal Mach Intell 28(11):1768–1783
9. Grau V, Noble, JA (2005) Adaptive multiscale ultrasound compounding using phase information. In: International conference on medical image computing and computer-assisted intervention, pp. 589–596. Springer
10. Hennersperger C, Baust M, Mateus D, Navab N (2015) Computational sonography. In: International conference on medical image computing and computer-assisted intervention, pp 459–466. Springer
11. Hennersperger C, Mateus D, Baust M, Navab N (2014) A quadratic energy minimization framework for signal loss estimation from arbitrarily sampled ultrasound data. In: International conference

on medical image computing and computer-assisted intervention, pp 373–380. Springer

12. Hindi A, Peterson C, Barr RG (2013) Artifacts in diagnostic ultrasound. Rep Med Imaging 6:29–48

13. Hung ALY, Chen E, Galeotti J (2020) Weakly-and semi-supervised probabilistic segmentation and quantification of ultrasound needle-reverberation artifacts to allow better ai understanding of tissue beneath needles. arXiv preprint arXiv:2011.11958

14. Hung ALY, Chen W, Galeotti J (2020) Ultrasound confidence maps of intensity and structure based on directed acyclic graphs and artifact models. arXiv preprint arXiv:2011.11956

15. Jensen JA (1999) Linear description of ultrasound imaging systems: notes for the international summer school on advanced ultrasound imaging at the Technical University of Denmark

16. Karamalis A, Wein W, Klein T, Navab N (2012) Ultrasound confidence maps using random walks. Med Image Anal 16(6):1101–1112

17. Kremkau FW, Taylor K (1986) Artifacts in ultrasound imaging. J Ultrasound Med 5(4):227–237

18. Lasso A, Heffter T, Rankin A, Pinter C, Ungi T, Fichtinger G (2014) Plus: open-source toolkit for ultrasound-guided intervention systems. IEEE Trans Biomed Eng 61(10):2527–2537

19. Meng Q, Sinclair M, Zimmer V, Hou B, Rajchl M, Toussaint N, Oktay O, Schlemper J, Gomez A, Housden J, Matthew J, Rueckert D, Schnabel J, Kainz B (2019) Weakly supervised estimation of shadow confidence maps in fetal ultrasound imaging. IEEE Trans Med Imaging 38(12):2755–2767

20. Mohebali J, Patel VI, Romero JM, Hannon KM, Jaff MR, Cambria RP, LaMuraglia GM (2015) Acoustic shadowing impairs accurate characterization of stenosis in carotid ultrasound examinations. J Vasc Surg 62(5):1236–1244

21. Mozaffari MH, Lee WS (2017) Freehand 3-D ultrasound imaging: a systematic review. Ultrasound Med Biol 43(10):2099–2124

22. Otsu N (1979) A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybernet 9(1):62–66

23. Quien MM, Saric M (2018) Ultrasound imaging artifacts: How to recognize them and how to avoid them. Echocardiography 35(9):1388–1401

24. Reusz G, Sarkany P, Gal J, Csomos A (2014) Needle-related ultrasound artifacts and their importance in anaesthetic practice. Br J Anaesth 112(5):794–802

25. Roche A, Pennec X, Malandain G, Ayache N (2001) Rigid registration of 3-D ultrasound with MR images: a new approach combining intensity and gradient information. IEEE Trans Med Imaging 20(10):1038–1049

26. Rohling R, Gee A, Berman L (1997) Three-dimensional spatial compounding of ultrasound images. Med Image Anal 1(3):177–193

27. Rohling R, Gee A, Berman L (1999) A comparison of freehand three-dimensional ultrasound reconstruction techniques. Med Image Anal 3(4):339–359

28. Rohling RN (1999) 3D freehand ultrasound: reconstruction and spatial compounding. Ph.D. thesis, Citeseer

29. Soldati G, Demi M, Smargiassi A, Inchingolo R, Demi L (2019) The role of ultrasound lung artifacts in the diagnosis of respiratory diseases. Expert Rev Respir Med 13(2):163–172

30. Tay PC, Acton ST, Hossack J (2006) A transform method to remove ultrasound artifacts. In: 2006 IEEE Southwest symposium on image analysis and interpretation, pp 110–114. IEEE

31. Tay PC, Acton ST, Hossack JA (2011) A wavelet thresholding method to reduce ultrasound artifacts. Comput Med Imaging Graph 35(1):42–50

32. Toet A (1989) Image fusion by a ration of low-pass pyramid. Pattern Recogn Lett 9(4):245–253

33. Trobaugh JW, Trobaugh DJ, Richard WD (1994) Three-dimensional imaging with stereotactic ultrasonography. Comput Med Imaging Graph 18(5):315–323

34. Virga S, Göbl R, Baust M, Navab N, Hennersperger C (2018) Use the force: deformation correction in robotic 3D ultrasound. Int J Comput Assist Radiol Surg 13(5):619–627

35. Win KK, Wang J, Zhang C, Yang R (2010) Identification and removal of reverberation in ultrasound imaging. In: 2010 5th IEEE conference on industrial electronics and applications, pp 1675–1680. IEEE

36. Xu X, Zhou Y, Cheng X, Song E, Li G (2012) Ultrasound intima-media segmentation using hough transform and dual snake model. Comput Med Imaging Graph 36(3):248–258

37. Zevallos N, Harber E, Abhimanyu, Patel K, Gu Y, Sladick K, Guyette F, Weiss L, Pinsky MR, Gomez H, Galeotti J, Choset H (2021) Toward robotically automated femoral vascular access

38. Zou KH, Warfield SK, Bharatha A, Tempany CM, Kaus MR, Haker SJ, Wells WM III, Jolesz FA, Kikinis R (2004) Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports. Acad Radiol 11(2):178–189