Real-time Spatial Compounding with Warping

Adrian R. Groves and Robert N. Rohling

Department of Electrical and Computer Engineering, University of British Columbia,
Vancouver, Canada

ABSTRACT

Spatial compound imaging via beam-steering aims to improve image quality through signal averaging. However, compounding techniques are vulnerable to speed-of-sound and refraction distortions in non-homogeneous tissue. We have developed a system to perform image-based non-rigid registration in real time. The goal is to increase image quality by improving the alignment of the ultrasound frames before compounding. Frames are acquired by a PC-based ultrasound machine (Ultrasonix Medical Corp, Vancouver, Canada), and transmitted to a Windows-based workstation through a high-speed network. Robust image-to-image registration (warping) is performed using block-based estimation of local shifts and thin-plate spline interpolation of shift vectors. Compound images are computed as a rolling average of the nine most recent warped frames. The process runs at 20 frames per second on a dual-processor Xeon workstation, demonstrating the feasibility of sophisticated real-time image processing on a standard PC platform. High speed is achieved through algorithm refinements, approximations in speed-critical sections, and low-level optimizations. The result is a fully-automatic real-time spatial compounding system with a demonstrated improvement in image quality. Tests of registration accuracy were performed on simulated data with realistic speckle patterns, using a 10% speed-of-sound variation and an 8° beam-steering angle. The average misalignment across the image was reduced by 70%, from 0.22 mm to 0.07 mm; in the deepest parts of the image, alignment was improved by 91%. Improved quality is demonstrated on images of a human forearm, which show visibly improved edge sharpness. This work is one of a series of projects demonstrating the ability of a new open-architecture ultrasound system.

Keywords: Spatial compounding, ultrasound, realtime, non-rigid registration, warping, radial basis function, thin-plate spline, PC architecture

1. INTRODUCTION

Real-time spatial compound imaging aims to improve image quality by averaging several coplanar ultrasound frames into a single image. These component frames are acquired in rapid succession, using electronic beam-steering to obtain frames from a number of different look angles. The displayed image is updated in real time as freshly-acquired frames replace old frames with the same beam-steering angle. Tests with both experimental and commercial spatial compounding systems have shown that spatial compounding is beneficial to many clinical applications, such as the diagnosis of breast anomalies^{1,2} and the examination of carotid plaques.^{3,4}

However, spatially compounded images can suffer from blurring due to misalignment between frames. The speed of sound in soft tissue can vary by as much as 14% (Table 1), introducing distortions due to refraction and speed-of-sound errors. In conventional (single-frame) imagery these slight shifts go unnoticed, but when multiple beam angles are used the beam-steered frames are each distorted differently and alignment errors result. We believe the blurring due to image misalignment causes a significant reduction in the sharpness of compound images. This difficulty has been previously noted by other researchers in 2D and 3D compounding.^{3,5} In 3D spatial compounding, misalignment is a more serious issue because of tissue movement during the relatively slower acquisition process, and image-to-image registration has been shown to significantly improve the quality of 3D ultrasound volumes.^{6,7}

We propose to develop a 2D spatial compounding system using image-to-image registration to realign the frames prior to compounding. Each frame is registered to the reference frame (without beam-steering) before

---

Further author information: (Send correspondence to R.N.R.)

R.N.R.: E-mail: rohling@ece.ubc.ca, Telephone: 1 604 822 2045, Fax: 1 604 822 5949

A.R.G.: E-mail: argroves@interchange.ubc.ca

Table 1. The speed of sound in selected soft tissues, with the percent speed-of-sound distortion (values are approximate; extracted from Hill,⁸ p.212). 1540 m/s is the standard estimate used to form ultrasound images.

Tissue	Speed of Sound	Variation from 1540 m/s
Muscle	1560–1630 m/s	1–6%
Breast	1420–1560 m/s	0–8%
Skin	1500–1720 m/s	0–12%
Fat	1350–1460 m/s	5–12%
Tendon	1750 m/s	14%

being added to the compound image. The registration should be performed in real time to preserve the interactive nature of an ultrasound exam. Speed-of-sound errors can vary throughout the material, so non-rigid registration (referred to as “warping”) is used to accurately describe the image distortions. To construct such a system, a development framework is needed which offers realtime digital image acquisition and software control over imaging parameters, such as the beam-steering angle.

The authors have already created such a framework called the Programmable Ultrasound Platform and Interface Library (PUPIL). PUPIL was developed as an open ultrasound platform to allow researchers a level of access and control that is not generally available in medical ultrasound systems. It is built upon the Ultrasonix 500RP (Ultrasonix Medical Corp., Burnaby, Canada): a commercial PC-based ultrasound system with an additional research package. PUPIL provides realtime access to image data at all stages of processing, and enables the imaging parameters to be dynamically reconfigured by software. It also provides a real-time interface to the 500RP from a high-performance workstation through a fiber-optic network. A brief description of this system can be found in Rohling *et al.*,⁹ with more details available in the technical report.¹⁰

We will demonstrate a spatial compounding system with warping that achieves real-time performance on standard PC hardware and produces compound images with improved quality. We will evaluate the improvement in frame alignment using simulated images, and show the improvement in image quality using sample images acquired from a human forearm.

2. METHOD

2.1. Hardware and development tools

Ultrasound images are acquired from the Ultrasonix 500RP using a 38mm linear 4–9 MHz probe and a depth setting of 30mm. The post-scan-converted frames are received from the 500RP as a series of gray-scale bitmaps, with each beam-steered image filling a trapezoidal region of its bitmap. The bitmap resolution at these settings is 13.8 pixels/mm. Warping and compounding computations are performed on a dual-processor workstation with two 2.8 GHz Intel Xeon processors and 1 GB of RAM.

The warping and compounding software was developed in C++ using Microsoft Visual Studio 6.0 and the Intel optimizing compiler (Intel C++ Compiler 7.0 for Windows). Standard multiprocessing directives (OpenMP) were used to split the workload between the two processors. In addition, the Intel Integrated Performance Primitives library (IPP) provided high-speed implementations of several image-processing routines.

Many modern processors have the ability to perform single-instruction multiple-data (SIMD) operations, which can greatly accelerate repetitive calculations. However, taking full advantage of these capabilities often involves restructuring code and accessing special instructions, introducing additional complexity. The Xeon’s Streaming SIMD Extensions 2 (SSE2) enable the processor to perform up to 16 integer operations in each cycle, and can be accessed by embedding inline assembly code in C++ functions. This low-level optimization is used sparingly in our compounding system. Four functions were optimized using short sections (approximately 35 instructions each) of assembly code.

2.2. Image acquisition with PUPIL

For the current stage of research, images are acquired and stored using the PUPIL interface to manually select beam angle and save images to disk. In each test case, the probe is clamped in place to eliminate any movement between the acquisition of the frames.

Nine frames at different transmit angles are used for each compound image, which is comparable to the number used in commercial systems and research systems.^{2,3} On our machine, we found that image quality degraded significantly at beam angles beyond $\pm 10^\circ$, so we used the eight beam-steering angles $\pm 2^\circ$, $\pm 4^\circ$, $\pm 6^\circ$, $\pm 8^\circ$ and the reference frame (0°) for all compound images.

2.3. Non-rigid image-to-image registration (warping)

The key feature of our compounding system is the introduction of a warping step to improve the alignment of tissue structures before the frames are compounded. The reference frame (without beam-steering) is used as the “gold standard,” and all beam-steered frames are registered to match it. A block-matching algorithm estimates the optimal alignment in different areas of the image. A grid of evenly-spaced target locations is selected in the reference frame, and a fixed-sized block of pixels around each target is extracted. In the beam-steered frame, all possible source locations within a predefined search range around the target location are considered. The best-matching source is selected by comparing the block of pixels surrounding each source to the target block. This best match gives a shift vector for each target location, and these are interpolated to provide a smooth field of vectors covering the entire frame. Re-mapping the pixels in the beam-steered frame according to the vector field yields a new “warped” frame. This is a slightly reshaped copy of the beam-steered frame, with its features better aligned relative to the reference frame. Since the shift vectors are interpolated from local estimates, the warping is non-rigid and can accurately describe a combination of shears, shifts, and rotations within the frame.

2.3.1. Block-based shift vector estimation

The choice of block size was motivated by the expected feature density in our ultrasound images. We found that a block size of 96 by 96 pixels (approx. 7 by 7 mm) ensured that nearly every block contained some significant features. An evenly-spaced 5-by-5 grid of target locations in the reference frame was used to make sure that every pixel in the frame contributed to the warping process. The search range of source locations was defined by a maximum shift of ± 10 pixels (± 0.73 mm) horizontally and ± 4 pixels (± 0.29 mm) vertically, which was determined experimentally to encompass any reasonably expected misregistration. It is assumed that the movement of each block of pixels between the two frames can be reasonably approximated by a translation, which is generally true when the rotations and shears are small relative to the block size.

At each target location vector that is tested, the measure of alignment is defined as the statistical correlation between the source and target blocks. Briefly, this is defined as the covariance between the two blocks, divided by the geometric mean of their individual variances.

In more detail, let A and B represent the reference and beam-steered blocks respectively, so that for a given shift vector $(\Delta x, \Delta y)$, the pixel $A(x, y)$ corresponds to $B(x + \Delta x, y + \Delta y)$. These pairs of corresponding pixels are arranged into vectors A_i and B_i . Note that because of the trapezoidal shape of the beam-steered frame and the relative block motion, some of the pixels of B may lie outside of the ultrasound data. Let I denote the set of all indices where B_i contains valid ultrasound data. Let μ_A and μ_B be the averages of A_i and B_i over all $i \in I$. Then the correlation coefficient C is defined by

$$C = \frac{\sum_{i \in I} (A_i - \mu_A)(B_i - \mu_B)}{\left(\sum_{i \in I} (A_i - \mu_A)^2 \cdot \sum_{i \in I} (B_i - \mu_B)^2 \right)^{\frac{1}{2}}}, \quad (1)$$

which is a value between +1 and -1, where +1 indicates a perfect match. The statistical correlation is independent of average image intensity and contrast, so it is a suitable measure to use with beam-steered frames where these averages sometimes varied from the reference frame.

This computation is repeated for each allowable vector within the shift range to build a cross-correlation matrix, with the peak (maximum value) indicating the best-matching shift vector. A single, clearly identifiable peak is typically present in the matrix, corresponding to the global maximum.

To accelerate the computation, a coarse-to-fine approach is used, with the coarse step using a half-scale image to search a broad shift range (equivalent to ± 10 pixels horizontally, ± 4 vertically in the full-sized image). The fine step searched a smaller shift range (± 2 pixels in each direction) centered on the result from the coarse step. It was found experimentally that this approach gave identical results to a single-stage approach in virtually all tests; slightly different results were occasionally obtained in regions with no visible features.

For additional speed gains, the innermost loop of the correlation step was rewritten in assembly to use the SSE2 instructions. Using SSE2 roughly doubles the speed of the calculations.

2.3.2. Smooth interpolation of motion vectors

After finding the shift vector at each target location, the next step is to construct a smooth interpolating function to give a shift vector at every point in the image. There are many suitable interpolation methods. We use the radial basis function (RBF) technique, which is used widely in biomedical applications^{11,12} and offers both flexibility and tunability (for example, it can provide a smooth fit while removing outliers).

Given N data points located at $x_1, x_2, \dots, x_N$, let $R_i(x) = \|x - x_i\|$ be the Euclidean distance from x to the i^{th} data point, and $\lambda_1 \dots \lambda_N$ be real constants. Then the general RBF interpolating function from $s(x) : \mathbb{R}^2 \rightarrow \mathbb{R}$ takes the form

$$s(x) = p(x) + \sum_{i=1}^N \lambda_i \phi(R_i(x)), \quad (2)$$

where $\phi(r)$ is a fixed real function, and $p(x)$ is the trend function (a low order polynomial with M monomial terms $p_1(x) \dots p_M(x)$). The solution is found by supplying the given data points $f(x_i)$ in

$$s(x_i) = f(x_i) \text{ for all } i = 1 \dots N \quad (3)$$

and applying the side conditions

$$\sum_{i=1}^N \lambda_i p_j(x_i) = 0 \text{ for all } j = 1 \dots M. \quad (4)$$

When applying the RBF method to registration, the horizontal and vertical components of the shift vectors are interpolated separately. A commonly used RBF for 2D registration is the thin-plate spline,¹² which uses the fundamental solution of the biharmonic equation ($\phi(r) = r^2 \log r$) and an affine trend function ($p(x) = a_0 + a_1 x_1 + a_2 x_2$). By defining the “smoothness” of a function in terms of minimizing the curvature (second derivatives), it can be shown that the thin-plate spline is the smoothest possible function passing through all the data points. In addition, shift vectors near the image boundaries are found by 0^{th} -order extrapolation from the nearest interpolated value.

The coefficients $\lambda_1 \dots \lambda_N$ and a_0, a_1, a_2 are computed by the inversion of an $(N+3) \times (N+3)$ matrix, which takes less than 0.1 ms for number of data points ($N = 5 \times 5$) in the interpolation. However, evaluating $s(x)$ directly at every pixel requires a much larger amount of computation — over 750 ms for each frame. Since the thin-plate-spline function has minimal curvature, we accelerate this step by sampling the RBF at regular intervals and using linear interpolation to evaluate the pixels between. One sample was taken in every 16 by 16 pixel (1.2 mm by 1.2 mm) square, yielding an average accuracy of 0.02 pixels and a maximum observed error of 0.4 pixels when compared to full samples of the RBF. This approximation reduced the interpolation time to 12 ms, and further optimization with SSE2 instructions cut this time in half.

2.3.3. Image remapping

Once the shift vectors have been calculated for every pixel in the frame, a new warped frame can be created from the original off-axis frame. The motion vectors indicate where each pixel in the warped frame should come from the original frame, and nearest-neighbour interpolation is used to create the warped frame. This remapping is performed by an IPP library routine, which is highly optimized for the Xeon processor.

2.4. Compounding and display

For our purposes, compounding the images means finding the average intensity of each pixel over all of the warped images and the reference image. The degree of compounding varies across the image because of the variable amount of overlap, so the average is computed on a pixel-by-pixel basis.

The nine component images are stored, and an accumulator keeps track of the sum of all the images. When a new warped image becomes available, it is added to the accumulator and the oldest image is subtracted and discarded. The accumulator also keeps an array of how many frames have contributed intensity values to each pixel, so the most current compound image is the quotient of the accumulator and the frame-count array. Any post-processing algorithms, such as contrast enhancement, can be applied to the final image, but in our implementation it is simply displayed or saved to disk.

2.5. Experimental verification

Our expectation is that the warping step improves the alignment of features, so that the compounding with warping produces a significantly sharper image than compounding without warping. The registration accuracy of the warping algorithm will be evaluated using simulation images, and sample images acquired from a human forearm will be processed by three different methods (no compounding, compounding without warping, and compounding with warping) to demonstrate the effectiveness of this technique in real tissue. Sample frames from the experimental samples are shown in Fig. 1.

2.5.1. Simulation

A pair of simulated frames was created solely to evaluate the registration accuracy of the warping algorithm. Simulations are ideally suited to this task because they offer great control over image content and precise knowledge of the shift vectors across the image.

Simulated ultrasound images were created using the Field II software program,¹³ which creates an ultrasound frame from a given scatterer distribution. The parameters and configuration were chosen to closely model the probe used in this paper. Field II uses a far-field approximation, so to obtain clear images the frames were extracted from depths of 30 mm to 65 mm from the probe. A layered checkerboard template was created in this region and populated with randomly distributed scatterers. Using the simulation program ensures that the speckle patterns in each frame are realistic.

Using our own ray-tracing model, we calculated the apparent shift of the scatterers between the 0° and 8° beam-steering angle, based on a three-layer phantom with a 10% speed of sound variation. The scatterer

Figure 1. The non-beam-steered images from the test data sets: (a) simulation and (b) human forearm.

distribution is shifted according to this apparent shift and the simulation repeated. Because the speckle patterns become decorrelated as the beam angle increases, we assumed the worse case and used completely decorrelated scatterer distributions for the two frames. This ensures that the warping algorithm is matching features, not speckle patterns. When warping is applied to the image pair, the estimated vector field can be compared to the actual applied shift, and the residual errors can be quantified.

2.5.2. Tests on a human subject

In our current configuration the beam-steering angle could only be changed at a slow rate, so the frames were acquired over a period of seconds rather than milliseconds. The probe was clamped and movement of the arm was minimized. This is done to demonstrate the ability of the proposed warping algorithm to correct for speed of sound and refraction errors, not probe motion between frames. For future configurations where the beam-steering angle is changed at a fast rate, probe-motion errors are expected to be small. The 0° frame from the acquired data set is shown in Fig. 1(b).

3. RESULTS

3.1. Real-time performance

Timing results are averages from a test data set containing a large number of images to simulate steady-state operation. Timing is measured for each section of code. On average, the computation of each image took 40.2 ms, broken down by section (See Table 2). Most of the computation time was spent performing the cross-correlation (20.0 ms) and the interpolation (6.9 ms).

A complete clinical implementation will also require the time to receive each bitmap from the 500RP and to draw the final compound image on the screen. Fung and Rohling¹⁰ measured these times for the PUPIL hardware at 8.4 ms and 7.1 ms respectively; furthermore, they determined that these steps can be performed concurrently. It is therefore reasonable to conclude that the average computation time will be under 50 ms per image, yielding a realtime frame rate of 20 fps.

3.2. Registration accuracy

The simulation images were used to test the accuracy of the warping algorithm. Known motion vectors were used to create the distortion in the simulated image, so the residual errors are easily computed across the entire image. Fig. 2 shows the estimated vector field used to realign the 8° frame, and the residual errors after registration. The initial misalignment had an average magnitude of 0.22 mm across the image, but warping reduced this by 70% to 0.07 mm (0.9 pixels). On the lower section of the image, the misalignment dropped from 0.44 mm to 0.04 mm (0.5 pixels), a 91% improvement.

Interestingly, absolute registration accuracy in the simulated image was better in the bottom part of the image (0.04 mm) than in the top part (0.10 mm). In fact, the two images are initially aligned perfectly in the top section, and warping introduces this small distortion. A close inspection of the two simulated images shows

Table 2. Computation time for the various steps of the spatial compounding algorithm. “Miscellaneous” includes various steps such as image copying and resizing.

Computation Step	Duration
Shift Estimates: Coarse	8.0 ms
Shift Estimates: Fine	12.0 ms
Interpolation	6.9 ms
Frame Remapping	3.9 ms
Compounding	3.1 ms
Miscellaneous	6.3 ms
Total	40.2 ms

Figure 2. (a) The estimated distortion vectors used for warping the 8° simulation frame into alignment with the 0° frame. (b) Residual errors across the image. All vectors are magnified by a factor of 3 for better visibility.

that the boundaries are sharper at the bottom of the image than at the top, presumably because of the far-field approximation used by Field II. This simply indicates that warping performs better with sharper features. However, even when the features are visibly blurred in the original frames, the registration error is still only about 1 pixel.

3.3. Sample tissue images

Test data sets were acquired from a human forearm (Fig. 3) and compounded. These images show that speckle is visibly reduced by either compounding method, and that warping can visibly improve compound image sharpness in images of real tissue. Further clinical tests are required to verify the value of spatial compounding with warping for aiding diagnostics and interventions.

4. CONCLUSION

This paper has shown that real-time registration can improve frame alignment, and thereby improve image sharpness, when realistic speed-of-sound distortions are present in the imaged specimen. Furthermore, compounding with warping still achieves speckle reduction over conventional imagery. This technique was implemented on standard hardware to provide clinicians with an additional imaging mode with some of the benefits of both conventional imaging and standard compound imaging.

Further testing will focus on developing techniques to measure the sharpness of features in compound images, so that the image quality benefits of compounding with warping can be better evaluated. Future work will lead towards creating a real-time compounding system with warping that is suitable for routine clinical use. Extensive testing by experienced ultrasonographers will then be used to make conclusions about the benefits of this technique for a number of clinical applications.

ACKNOWLEDGMENTS

This work is supported by the Institute for Robotics and Intelligent Systems, the Canadian Foundation for Innovation, and the National Sciences and Engineering Research Council of Canada. The authors would also like to thank Wilson Fung and Pedram Lajevardi for their contributions to the preliminary warping and compounding code.

Figure 3. Images of the underside of a human forearm in (a) conventional imaging, (b) compound imaging without warping, (c) compound imaging with warping. (d) A plot of the estimated shift vectors used to warp the -8° frame.

REFERENCES

1. R. Entrekin, P. Jackson, J. R. Jago, and B. A. Porter, "Realtime spatial compound imaging in breast ultrasound: technology and early clinical experience," *MedicaMundi* **43**, pp. 35–43, 1999.
2. S. Huber, M. Wagner, M. Medl, H. Czembirek, "Real-time spatial compound imaging in breast ultrasound," *Ultrasound Med. Biol.* **28**, pp. 155–163, 2002.
3. S. K. Jespersen, J. E. Wilhjelm, and H. Sillesen, "In vitro spatial compound scanning for improved visualization of atherosclerosis," *Ultrasound Med. Biol.* **26**, pp. 1357–1362, 2000.
4. S. C. Kofoed, M. M. Grønholdt, J. E. Wilhjelm, J. Bismuch, and H. Sillesen, "Real-time spatial compound imaging improves reproducibility in the evaluation of atherosclerotic carotid plaques," *Ultrasound Med. Biol.* **27**, pp. 1311–1317, 2001.
5. J. F. Krücker, C. R. Meyer, G. L. LeCarpentier, J. B. Fowlkes, and P. L. Carson, "3D spatial compounding of ultrasound images using image-based nonrigid registration," *Ultrasound Med. Biol.* **26**, pp. 1475–1487, 2000.
6. R. N. Rohling, A. H. Gee, and L. Berman, "Automatic registration of 3-D ultrasound images," *Ultrasound Med. Biol.* **24**, pp. 841–854, 1998.
7. G. Xiao, J. M. Brady, and A. Noble, "Nonrigid registration of 3-D free-hand ultrasound images of the breast," *IEEE Trans. Med. Imag.* **11**, pp. 405–412, 2002.
8. C. R. Hill, *Physical principles of medical ultrasonics*, John Wiley & Sons, Chichester, UK, 1986.
9. R. Rohling, W. Fung, and P. Lajevardi, "PUPIL: programmable ultrasound platform and interface library," in *Lecture notes in computer science* **2879**, R. Ellis and T. M. Peters, eds., Springer-Verlag, New York, NY, pp. 424–431, 2003.
10. W. Fung and R. Rohling, "Design of an open-architecture ultrasound acquisition system for real-time processing and control," Technical Report TR002, University of British Columbia, 2003. <http://www.ece.ubc.ca/~rohling/publications/tr002.pdf>.
11. J. C. Carr, R. Fright, and R. K. Beatson, "Surface interpolation with radial basis functions for medical imaging," *IEEE Trans. Med. Imag.* **16**, pp. 96–107, 1997.
12. F. L. Bookstein, "Principal warps: thin-plate splines and the decomposition of deformations," *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, pp. 567–585, 1989.
13. J. A. Jensen, "Field: a program for simulating ultrasound systems." Paper presented at the 10th Nordic-Baltic Conference on Biomedical Imaging Published in *Medical & Biological Engineering & Computing* **34**, Supplement 1, Part 1, pp. 351–353, 1996.