**Medical Image Analysis** (1996/7) volume 1, number 3, pp 177–193


# Three-dimensional spatial compounding of ultrasound images

Robert Rohling[1*], Andrew Gee[1] and Laurence Berman[2]

[1]Department of Engineering, Trumpington Street, Cambridge CB2 1PZ, UK
[2]Department of Radiology, Addenbrooke's Hospital, Cambridge CB2 2QQ, UK

**Abstract**
One of the most promising applications of 3-D ultrasound lies in the visualization and volume estimation of internal 3-D structures. Unfortunately, the quality of the ultrasound data can be severely degraded by artefacts and speckle, making automatic analysis of the 3-D data sets very difficult. In this paper we investigate the use of 3-D spatial compounding to reduce speckle. We develop a new statistical theory to predict the improvement in signal-to-noise ratio with increased levels of compounding, and verify the predictions empirically. We also investigate how registration errors can affect automatic volume estimation of structures within the compounded 3-D data set. Having established the need to correct these errors, we present a novel reconstruction algorithm which uses landmarks to register each B-scan accurately as it is inserted into the voxel array. In a series of *in vitro* and *in vivo* trials, we demonstrate that 3-D spatial compounding is very effective for improving the signal-to-noise ratio, but correction of registration errors is essential.

*Keywords:* 3-D ultrasound imaging, registration, spatial compounding, volume estimation

*Received November 4, 1996; revised January 8, 1997; accepted January 9, 1997*

## 1. INTRODUCTION

Diagnostic ultrasound imaging has been in routine medical use for more than 30 years. Conventional diagnostic imaging is performed with a hand-held probe which transmits ultrasound pulses into the body and receives the echos. The magnitude and timing of the echos are used to create a 2-D grey-level image (B-scan) of a cross-section of the body in the scan plane. One of the limitations of conventional imaging is the requirement that the physician mentally reconstruct the 3-D anatomy given multiple 2-D slices. Research is underway to overcome this limitation using 3-D ultrasound imaging.

High-quality, instantaneous 3-D imaging remains a long-term goal of medical ultrasound research. One promising approach centres around the development of a new type of phased array probe, which sends and receives echos from a 2-D array of elements (instead of the usual 1-D array), but several technical challenges must be overcome before such probes receive clinical acceptance (Smith *et al.*, 1995). Alternative approaches, which make use of conventional 2-D ultrasound technology, include the free-hand and swept volume techniques (Rankin *et al.*, 1993; Steiner *et al.*, 1994). Instead of taking an instantaneous 3-D snapshot, these techniques construct a 3-D volume from a number of 2-D B-scans acquired in rapid succession.

In the free-hand paradigm, a 3-D position sensor is attached to the probe, so that each B-scan can be labelled with the position and orientation of the scan plane (see Figure 1). Subsequent processing can build up a 3-D description of the imaged anatomy, in much the same manner as is possible with CT or MRI, but with less expensive or invasive technology. Swept-volume systems also use a set of B-scans to create a 3-D description, but require a specially constructed probe to sweep the B-scans mechanically through a volume of interest. In this paper we focus on the free-hand technique.

Physicians have indicated that there is significant utility in 3-D ultrasound imaging of a variety of anatomical structures, including the foetus (Chervenak *et al.*, 1993), vascular structure (Franseschi *et al.*, 1992), gall bladder (Fine *et al.*, 1991), breast (Moskalik *et al.*, 1995), kidney (Gilja *et al.*, 1995), eye (Downey *et al.*, 1996) and heart (Salustri and Roelandt, 1995). In review articles about 3-D ultrasound (Hottier and Collet Billon, 1990; Rankin *et al.*, 1993), the authors suggest that 3-D visualization and volume estimation are its

*Corresponding author
(e-mail: rnr20@eng.cam.ac.uk)

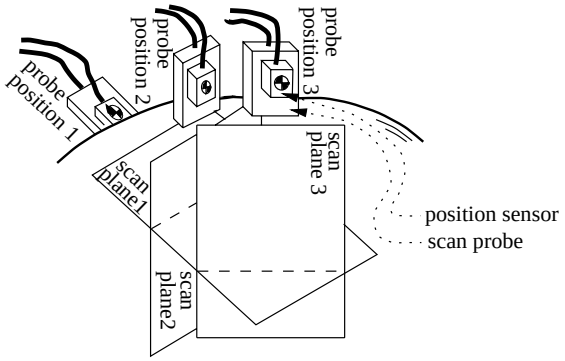



**Figure 1.** 3-D free-hand ultrasound imaging. Free-hand imaging allows the physician to move the probe as in a normal ultrasound examination. The position sensor measures the position and orientation of each scan plane. Note that the planes intersect each other.

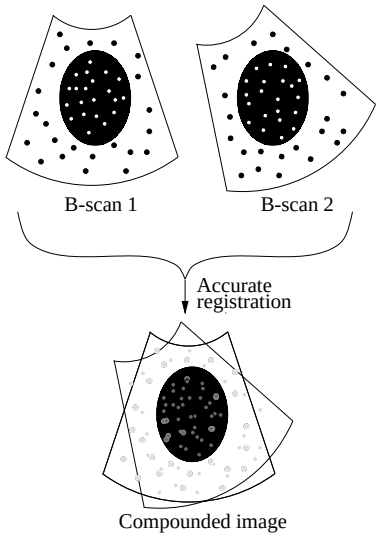



**Figure 2.** Spatial compounding. This is a simple illustration of 2-D spatial compounding. Two scans of the same plane are accurately registered and then averaged together to produce a compounded image with an improved SNR. The principle extends to 3-D, where compounding can be performed wherever scan planes intersect, see Figure 1.

most attractive capabilities. The main difficulty in performing these tasks automatically is that the quality of the ultrasound data can be severely degraded by artefacts and speckle. Speckle is common to all ultrasound images, a product of the constructive–destructive interference of the ultrasound echos. While trained ultrasonographers can occasionally infer diagnostically useful information from speckle patterns, speckle can inhibit automatic analysis of ultrasound data. For this reason there has been considerable research into speckle reduction.

There are three main techniques for reducing speckle in ultrasound images (Evans and Nixon, 1995): filtering (Crawford *et al.*, 1993; Evans and Nixon, 1995; Sakas *et al.*, 1995), phase-based methods (Leeman and Seggie, 1987) and compounding. Compounding can be performed spatially (Kerr *et al.*, 1986; Shankar, 1986; Trahey *et al.*, 1986b; Hernandez *et al.*, 1996) or with multiple frequencies (Trahey *et al.*, 1986a). While each technique has its advantages and disadvantages, spatial compounding is particularly attractive with 3-D ultrasound, since it can be performed with a standard free-hand system without the need for any internal modifications.

The principle behind spatial compounding is to image the region of interest repeatedly, from different look directions and then average the values from the intersecting B-scans when constructing the 3-D data set (see Figure 2). The speckle signal, which de-correlates from different look directions, is suppressed by the averaging operation. Conversely, real anatomical features (tissue boundaries, for example) will be observed in the same location from all look directions. Provided the registration of the scan planes is accurate, the averaging operation will highlight the real anatomical features.

There has been considerable research in the past on the spatial compounding of multiple B-scans for 2-D image quality improvement. We define the signal-to-noise ratio (SNR) as the ratio of the mean grey level to the standard deviation for an image with no resolvable structures[a]. It has been established (Burckhardt, 1978; Kerr *et al.*, 1986; Hernandez *et al.*, 1996) that compounding multiple B-scans lying in the same plane can improve the SNR by as much as the square root of the number of B-scans used.

The main drawback of spatial compounding is that it requires accurate registration: registration errors will place the same anatomical feature seen from different look directions at different positions in the reconstructed volume. This phenomenon can be seen in Figure 12, which shows a reduction in speckle as more B-scans are averaged, accompanied by a blurring of the imaged boundaries. While the largest sources of registration error are likely to be from inaccurate B-scan position measurement and tissue motion during the scan (Rohling and Gee, 1996), refraction of the ultrasound beam and other imaging effects also contribute.

In this paper we take a fresh look at spatial compounding, this time from the perspective of 3-D free-hand ultrasound. Though our main focus is on speckle reduction, spatial compounding can also improve image quality by reducing several other kinds of artefacts, such as shadowing. Every free-hand system has to deal with compounding in some manner, since

[a]This is a commonly used figure of merit for imaging systems.

it is almost inevitable that the scan planes will intersect. Here we propose deliberate, extensive compounding, with the aim of producing high-quality 3-D data sets which lend themselves to automatic segmentation for visualization and volume measurement. The key to effective spatial compounding is to achieve a sufficiently high registration accuracy. Relying on the position sensor alone is usually not sufficient: there may be small errors in its calibration and it does not take into account motion of the target or within-plane imaging artefacts. It is therefore necessary to improve the registration using image-based techniques.

Little work on the registration of 3-D ultrasound data sets is evident in the literature. In one exception (Moskalik *et al.*, 1995), two separate data sets were retrospectively registered using manual landmark matching. This constitutes a labour-intensive solution to a specific registration problem. Here we propose an automatic, incremental registration algorithm for use with generic free-hand ultrasound imaging. Similarly, little work on 3-D ultrasound compounding has been performed. Only one brief article (Nelson and Pretorius, 1994) has cited the improvements possible by 3-D compounding, but it simply stated the need for accurate registration without providing further detail.

In this paper we tackle three major objectives. The first is to demonstrate how accurate 3-D registration can be achieved. We describe a technique that takes sensor-based measurements of B-scan positions and applies small adjustments to align anatomical landmarks in the reconstructed volume. The second objective is to demonstrate how spatial compounding, coupled with accurate registration, can dramatically improve the SNR of the reconstructed data. Our final objective is to develop a statistical theory of 3-D spatial compounding and establish agreement between the observed and predicted improvements in SNR. In the course of the investigation we perform two empirical studies: an *in vitro* phantom study and an *in vivo* study. The phantom study allows the 3-D reconstruction process to be evaluated and verified before proceeding to the *in vivo* study.

## 2. ACQUISITION SYSTEM AND TEST SUBJECTS

The acquisition system comprises a Toshiba model SSA-270A/HG ultrasound scanner, a standard 2-D probe and a position sensor. The phantom study used a 7 MHz linear array probe and the *in vivo* study used a 3.75 MHz convex curvilinear array probe. The position and orientation of each scan plane, relative to a fixed transmitter, are measured by an AC magnetic field receiver (Polhemus FASTRAK) mounted on the probe. Images from the video output of the scanner are recorded by an 8-bit framegrabber at a rate of 5 frames/s. The images and the position data are stored in the memory of a Sun SparcStation 10 workstation.

Laboratory tests were first performed on a phantom comprising an egg-shaped latex balloon filled with a combination of water, ultrasound coupling gel and talcum powder. This type of phantom was used because B-scans of its cross-section produce images of a uniformly speckled interior, a sharp boundary and a uniformly speckled exterior with a lower mean grey level. Grey-level statistics can be easily measured for the interior and exterior regions. Figure 9a shows a typical B-scan of the phantom. The phantom was mounted in a bath of water at an elevated temperature so that the propagation speed of ultrasound in the water approximated the speed in human tissue. An *in vivo* examination was also performed on the gall bladder of a healthy human subject. The examination was performed by an experienced ultrasonographer in a manner fairly consistent with a normal ultrasound examination.

The manufacturer states that the RMS (root mean squared) accuracy of the FASTRAK position measurements is 0.8 mm with a resolution of 0.38 mm. The angle measurements are stated to have an RMS accuracy of $0.15^\circ$ with a resolution of $0.025^\circ$. The accuracy of imaging a point is determined by the accumulated sources of error at each stage of the data acquisition process, including the speed-of-sound estimate, the FASTRAK readings and the calibration of the FASTRAK receiver to the plane of the B-scan. The overall accuracy can be estimated from the RMS value of the registration errors corrected in the phantom study: the RMS position error was 3.5 mm and the RMS orientation error was $1.2^\circ$.

## 3. 3-D RECONSTRUCTION

### 3.1. Reconstruction without registration

Most 3-D free-hand systems use similar algorithms to construct a 3-D data set from the individual B-scans. A typical reconstruction algorithm (Trobaugh *et al.*, 1994) is illustrated in Figure 3, with a detailed description following below.

Each B-scan is represented as a 2-D array **P** of intensity values $p_{mn}$. The reconstruction volume takes the form of a 3-D voxel array (or cuberille) **C**. Each element $c_{ijk}$ of **C** represents a voxel in space. The voxel size is chosen *a priori*: small voxels (though no smaller than the pixel dimensions) produce high-resolution reconstructions, larger voxels produce lower resolution reconstructions. While high-resolution reconstructions reveal more detail, they also require considerable computational resources to generate and manipulate. There is a fundamental trade-off between ease of data manipulation and resolution.

Figure 4 depicts the four coordinate systems used for reconstruction. The position sensor measures the relative

1. acquire 2-D image **P** and associated position data ${}^{T}\mathbf{T}_{R}$
2. insert image **P** into reconstruction volume **C**
   - 2.1 determine location of pixel $p_{mn}$ with respect to **C**
     ${}^{C}\underline{\mathbf{x}} = {}^{C}\mathbf{T}_{T}\ {}^{T}\mathbf{T}_{R}\ {}^{R}\mathbf{T}_{P}\ {}^{P}\underline{\mathbf{x}} \longleftrightarrow {}^{C}\underline{\mathbf{x}} = \mathbf{T}\ {}^{P}\underline{\mathbf{x}}$
   - 2.2 if nearest voxel $c_{ijk}$ in **C** is empty, set to $p_{mn}$
   - 2.3 else set $c_{ijk}$ to weighted average of existing $c_{ijk}$ and $p_{mn}$
     $c_{ijk} = \frac{n \times c_{ijk}}{n+1} + \frac{p_{mn}}{n+1}$
     where $n$ is incremented after each calculation of $c_{ijk}$
3. repeat from step 1.

**Figure 3.** Reconstruction algorithm without registration.

position and orientation of the receiver with respect to the transmitter. These measurements are converted into a $4 \times 4$ homogeneous transformation matrix ${}^{T}\mathbf{T}_{R}$. A standard notation is used to describe ${}^{T}\mathbf{T}_{R}$ as the transformation from the coordinate system at the receiver (**R**) to the coordinate system at the transmitter (**T**).

The position of a pixel $p_{mn}$ with respect to its plane (**P**) is expressed as a homogeneous vector ${}^{P}\underline{\mathbf{x}}$. The pixel position, with respect to the cuberille coordinate system (**C**), can be determined by transformation to the receiver coordinate system, then to the transmitter and finally to the reconstruction volume via ${}^{R}\mathbf{T}_{P}$, ${}^{T}\mathbf{T}_{R}$ and ${}^{C}\mathbf{T}_{T}$ respectively. ${}^{R}\mathbf{T}_{P}$ describes the transformation between the corner of the scan plane and the coordinate system of the receiver. It is determined by a simple calibration method similar to the one described in Hughes *et al.* (1996), and remains constant throughout the reconstruction. ${}^{C}\mathbf{T}_{T}$ describes the transformation from the transmitter to the corner of the cuberille. It is set to the limits of the reconstruction volume that the physician scans and also remains constant throughout the reconstruction. The cumulative matrix multiplication of ${}^{C}\mathbf{T}_{T}\ {}^{T}\mathbf{T}_{R}\ {}^{R}\mathbf{T}_{P}$ is abbreviated to **T**.

Before the start of the examination, the voxels in the reconstruction volume are all set to zero. As each B-scan is acquired, each voxel $c_{ijk}$ is adjusted according to the pixels $p_{mn}$ which intersect it. A single voxel will envelop many pixels if the voxel size is larger than the B-scan pixel size. Each voxel may also be intersected again by future B-scans. These possibilities are dealt with by step 2.3 of the reconstruction algorithm, which describes a compounding operation to average all pixels that intersect a voxel.

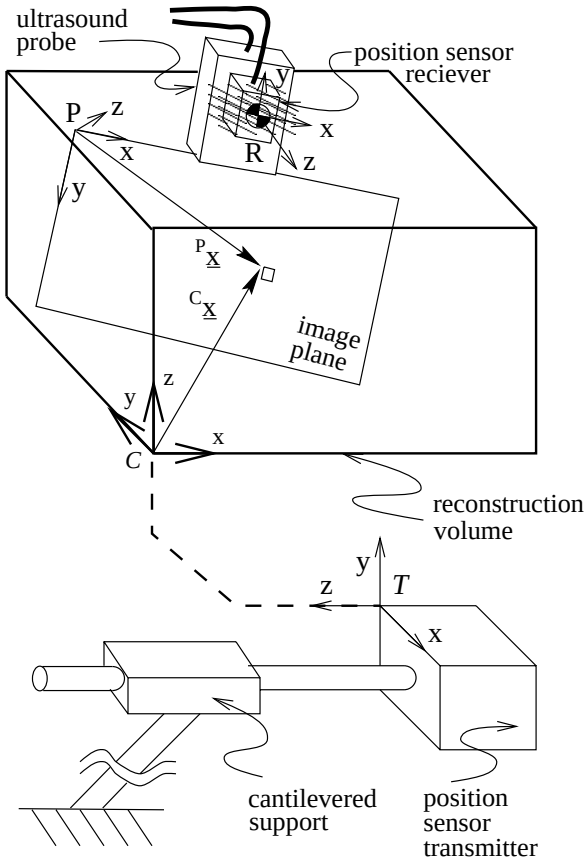



**Figure 4.** Coordinate systems used for reconstruction. The following notation is used to label coordinate systems: **T** = Transmitter, **C** = Cuberille, **P** = Plane and **R** = Receiver.

After a substantial portion of **C** is filled, it can be displayed on a computer monitor by several different methods, including volume rendering, surface rendering and any-plane slicing. Examples of any-plane slicing can be found in Figures 11, 12, 13, 16 and 17, while surface rendering is used in Figure 14. Hereafter, the term slice is used to indicate an image produced by any-plane slicing.

### 3.2. Reconstruction with registration

*3.2.1. Overview*

While the standard reconstruction algorithm is adequate for many tasks, it is not ideal when acquiring heavily compounded data sets, where voxels are intersected many times by B-scans acquired from a variety of look directions in the course of a relatively lengthy examination. Such reconstructions tend to be plagued by registration errors, caused by inaccurate B-scan position measurement, motion of the target and other imaging artefacts: see, for example, Figure 12. To

acquire high-quality, spatially compounded data sets, we have to correct these registration errors.

A variety of image registration techniques are available (Brown, 1992). Correlation-based techniques are the simplest but do not work well with 3-D ultrasound, since they are sensitive to artefacts such as shadows and reverberations. Changes in the B-scan image arising from changes in the ultrasound machine settings (such as the time–gain compensation curve) also affect correlation. Furthermore, correlation techniques are very inefficient when applied to 3-D (as opposed to 2-D) registration problems. For these reasons we prefer to attempt registration via landmarks.

Landmarks are anatomical features which are prominent in the B-scan images. When a newly acquired B-scan is compounded into a filled (or partially filled) voxel array, any landmarks in the B-scan should align with existing landmarks in the voxel array. Registration errors can be corrected by searching for corresponding landmarks and re-positioning the new B-scan plane so that the landmarks are brought closer together.

Registration has to be performed with respect to a reliable baseline. If all the B-scans are re-positioned, it is possible to construct a voxel array where all landmarks are in perfect alignment, but the reconstruction bears little resemblance to the underlying anatomy (for instance, the voxel array could contain a sheared image of the anatomy). For this reason, the ultrasound examination must commence with a quick pass over the region of interest, filling most of the voxels with little overlap of the B-scans. No attempt is made to register these initial B-scans: they act as the baseline. At this stage the 3-D data set will be noisy (due to speckle) but relatively free from registration errors[a]. Subsequent passes over the region of interest, from different look directions, are compounded into the voxel array to reduce the noise. The inevitable registration errors are automatically corrected by landmark alignment as each new B-scan is acquired.

Figure 5 describes a reconstruction algorithm with landmark-based registration. The main difference between this algorithm and the algorithm without registration (Figure 3) is that the image transformation $\mathbf{T}$ is replaced by $\mathbf{T}^*$, the optimal transformation that registers the landmarks in the image $\mathbf{P}$ with those already present in the voxel array $\mathbf{C}$. In this study, $\mathbf{T}^*$ is constrained to a six-degree-of-freedom rigid-body transformation, consistent with the expected sources of registration error (motion of the target and inaccurate B-scan position measurements).

[a] Any small registration errors that occur at this stage will not be corrected, but since the true shape of the anatomy is unknown, these initial B-scans provide the best estimate of the true shape.

1. acquire 2-D image $\mathbf{P}$ and associated position data ${}^{T}\mathbf{T}_{R}$
2. find landmarks $l_{mn}$ in image (see Figure 6)
3. if the scan plane does not intersect existing data (previous B-scans) in $\mathbf{C}$
   - 3.1 insert image $\mathbf{P}$ into reconstruction volume $\mathbf{C}$
     - 3.1.1 determine location of pixel $p_{mn}$ with respect to $\mathbf{C}$ ${}^{C}\underline{\mathbf{x}} = {}^{C}\mathbf{T}_{T}\,{}^{T}\mathbf{T}_{R}\,{}^{R}\mathbf{T}_{P}\,{}^{P}\underline{\mathbf{x}} \longleftrightarrow {}^{C}\underline{\mathbf{x}} = \mathbf{T}\,{}^{P}\underline{\mathbf{x}}$
     - 3.1.2 if nearest voxel $c_{ijk}$ in $\mathbf{C}$ is empty, set to $p_{mn}$
     - 3.1.3 else set $c_{ijk}$ to weighted average of existing $c_{ijk}$ and $p_{mn}$
   - 3.2 insert $l_{mn}$ into volume of landmarks $\mathbf{L}$
     - 3.2.1 determine location of $l_{mn}$ (at ${}^{P}\underline{\mathbf{x}}$) with respect to $\mathbf{C}$ ${}^{C}\underline{\mathbf{x}} = \mathbf{T}\,{}^{P}\underline{\mathbf{x}}$
     - 3.2.2 assign nearest vector in $\mathbf{L}$ to $l_{mn}$
   - 3.3 repeat from step 1.
4. if scan plane intersects existing data in $\mathbf{C}$, registration is performed
   - 4.1 optimize $\mathbf{T}$ to align $l_{mn}$ with landmarks in $\mathbf{L}$ $\mathbf{T} \longrightarrow \mathbf{T}^*$
   - 4.2 insert image $\mathbf{P}$ into reconstruction volume $\mathbf{C}$ via $\mathbf{T}^*$
     - 4.2.1 determine location of $p_{mn}$ with respect to $\mathbf{C}$ ${}^{C}\underline{\mathbf{x}} = \mathbf{T}^*\,{}^{P}\underline{\mathbf{x}}$
     - 4.2.2 if nearest voxel $c_{ijk}$ in $\mathbf{C}$ is empty, set to $p_{mn}$
     - 4.2.3 else set $c_{ijk}$ to weighted average of existing $c_{ijk}$ and $p_{mn}$
   - 4.3 repeat at step 1.

**Figure 5.** Reconstruction algorithm with registration.

### *3.2.2. Detection of landmarks*

The landmarks used in this study are edge elements (edgels) automatically extracted by the Canny edge detection algorithm (Canny, 1986). The edgels are produced at the resolution corresponding to the voxel size. The kernel size in the Canny convolution is given in Table 1.

The edgel set is further pruned by chaining neighbouring edgels together and eliminating short chains. Chains with fewer than three elements were eliminated because speckle typically produces one- or two-element chains, while organ boundaries produce longer chains: deleting short chains

**Table 1.** Registration algorithm parameters.

| Parameter | Organ phantom | *In vivo* study |
|---|---|---|
| Canny kernel size (mm) | 5 | 6 |
| Landmark search radius (mm) | 6 | 10 |
| Consensus set threshold (%) | 40 | 12 |

The size of the kernel used in the Canny convolution represents the best trade-off between the accuracy of boundary localization and sensitivity to speckle noise. The radius of the landmark search volume is related to the expected RMS registration error, which was estimated in Section 2. The higher value for the *in vivo* study reflects additional motion errors. The consensus set threshold is specified as a proportion of the total number of landmarks in each B-scan. Its value is based on an estimated ratio of the number of true landmarks to the total number of landmarks in a typical B-scan. The lower value for the *in vivo* study reflects the larger number of spurious landmarks detected in *in vivo* B-scans.

therefore reduces the number of false landmarks. Figure 6 illustrates the landmark detection procedure.

*3.2.3. Selection of correspondences*

As the B-scans are acquired, landmarks are stored in a 3-D vector array **L** which is aligned with **C**. When a B-scan intersects non-empty voxels in **C**, $\mathbf{T}^*$ is determined by finding correspondences between landmarks in the new B-scan and landmarks in **L**.

A minimum number of intersections between pixels in the scan plane and non-empty voxels in **C** is needed for accurate registration. In the phantom study, landmark registration was attempted only when more than 25% of the pixels in the B-scan intersected filled voxels in **C**. In practice, after the first sweep in the ultrasound examination, all subsequent B-scans have levels of intersections of almost 100%.

Potential correspondences between landmarks in the B-scan and landmarks in **L** are found by searching a spherical volume in **L** for each landmark in the B-scan. Since the B-scans are acquired in rapid succession, any registration error will vary slowly from one B-scan to the next: we can use this observation to place the spherical search volume at an appropriate location in the voxel array, and also limit its size (see Figure 7). We are effectively tracking the registration error, which is far more efficient than performing an unconstrained search for $\mathbf{T}^*$ for each B-scan. The size of the search volume is given in Table 1.

*3.2.4. Determination of transformation matrix and compounding*

The set of correspondences produces more constraints than are required to determine $\mathbf{T}^*$. For example, each B-scan

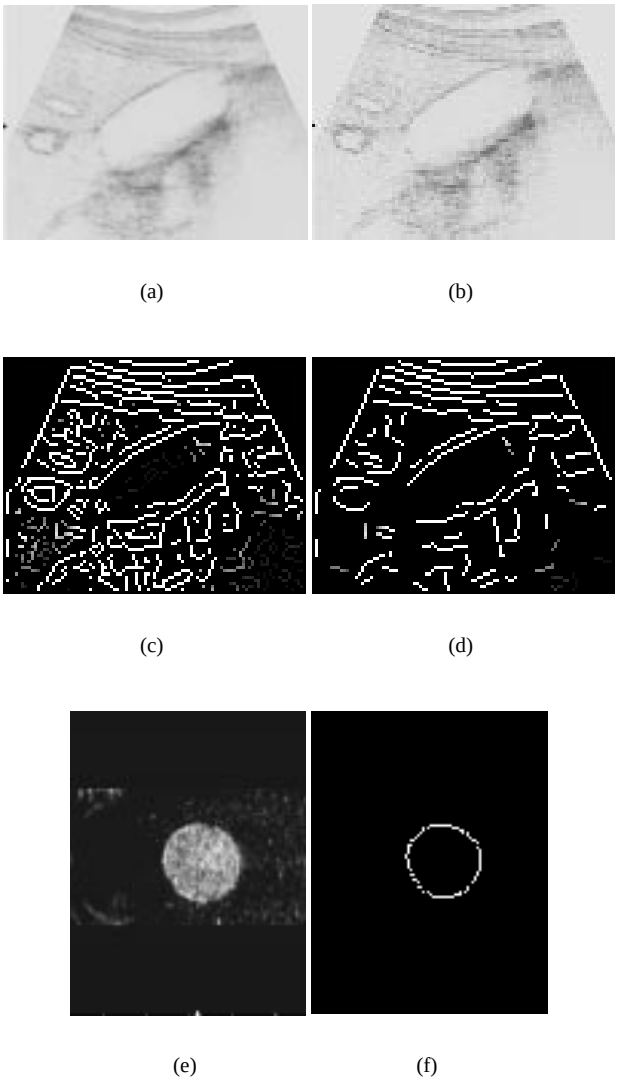



**Figure 6.** Landmark detection. Image (**a**) is the original high-resolution (490 × 380) image of a human gall bladder. Notice the speckling throughout the image, shadow-like artefacts and non-uniform intensity of the organ boundary. Image (**b**) is image (a) at the reduced resolution (99 × 77) required for the reconstruction volume. Image (**c**) depicts the 1409 edgels extracted from image (b). Image (**d**) depicts the chains that are formed from the edgels in (c). Notice the reduced number of edgels (854) in image (d)—the edgels forming small lines and circles are eliminated. As well as the organ boundary, the detector also finds edges which do not correspond to real physical structures: these must be tolerated by the registration algorithm. Images (**e**) and (**f**) show the same detection procedure applied to the phantom.

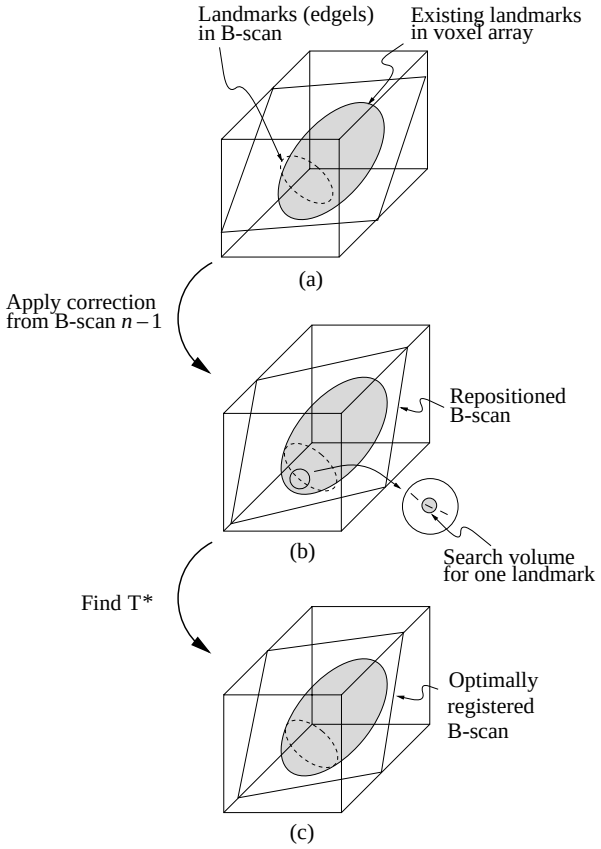
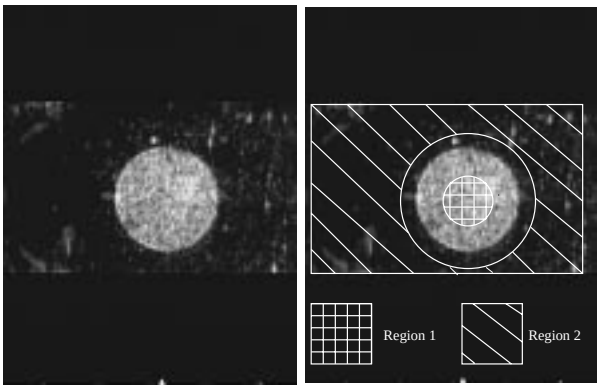



**Figure 7.** Landmark-based registration. In (**a**), B-scan $n$ is inserted into the voxel array at the location indicated by the position sensor. Significant registration errors are evident (exaggerated here for clarity). After applying the rigid-body correction found from registering B-scan $n-1$, only small residual errors remain (**b**). These are corrected by landmark-based registration. A small spherical search volume is defined around each re-positioned landmark (edgel) in **P**: any landmarks in **L** found within this volume are marked as candidate correspondences for the landmark in **P**. RANSAC regression is used to find the optimal rigid-body transformation $\mathbf{T}^*$ which aligns as many of the corresponding landmarks as possible (see Figure 8). Finally, the B-scan is inserted into its optimal position in the voxel array using $\mathbf{T}^*$ (**c**).

of the phantom generated ∼1200 candidate correspondences, but only three are needed to determine $\mathbf{T}^*$. A least-squares estimation of $\mathbf{T}^*$ is inappropriate, since many of the candidate correspondences are outliers. Instead, the RANSAC regression technique (Fischler and Bolles, 1981) is used to determine $\mathbf{T}^*$ and reject the erroneous correspondences. Details of RANSAC regression can be found in Figure 8. RANSAC is sufficiently robust to tolerate a significant proportion of erroneous correspondences.

1. Randomly pick three pairs of corresponding landmarks from the full set of correspondences.
2. Reject these correspondences if they are not consistent with a rigid-body transformation.
3. Reject these correspondences if the landmarks in **P** are too close together or collinear (otherwise the calculation of **T** in step 4 is ill-conditioned).
4. Calculate a linear affine transformation **T** which brings the three landmarks in **P** into precise registration with the corresponding landmarks in **L**. The pruning in step 2 ensures that **T** represents a rigid-body transformation.
5. Transform all the remaining landmarks in **P** by **T**.
6. Count how many of the transformed landmarks in **P** register with their counterparts in **L**. Those landmarks that do register contribute to the consensus set for **T**. The remaining landmarks are deemed outliers.
7. Repeat from 1 until a **T** is found with a consensus set larger than a preset threshold: this **T** becomes the estimate of the optimal transformation $\mathbf{T}^*$.

**Figure 8.** RANSAC regression for determining $\mathbf{T}^*$. The consensus set threshold is given in Table 1.




(a) Example B-scan (b) Regions

**Figure 9.** Regions of the phantom B-scans. To investigate the effect of compounding on the image statistics, two regions are defined: region 1 and region 2. Region 1 lies completely inside the latex balloon and region 2 on the outside. Similar volumetric regions are defined for the 3-D reconstruction volume.

**Table 2.** High-resolution reconstruction without registration.

| No. of | Region 1 | | | Region 2 | | |
|---|---|---|---|---|---|---|
| B-scans | mean | std dev. | SNR | mean | std dev. | SNR |
| 100 | 103.76 | 18.37 | 5.65 | 33.68 | 5.76 | 5.85 |
| 150 | 102.45 | 18.02 | 5.68 | 33.60 | 5.49 | 6.12 |
| 200 | 103.09 | 17.65 | 5.84 | 33.39 | 5.14 | 6.50 |
| 250 | 103.33 | 17.03 | 6.07 | 32.23 | 4.77 | 6.97 |
| 300 | 102.80 | 16.55 | 6.21 | 33.17 | 4.66 | 7.12 |
| 350 | 102.30 | 16.00 | 6.39 | 33.07 | 4.45 | 7.43 |
| 400 | 101.41 | 15.74 | 6.44 | 33.01 | 4.34 | 7.61 |

**Table 3.** High-resolution reconstruction with registration.

| No. of | Region 1 | | | Region 2 | | |
|---|---|---|---|---|---|---|
| B-scans | mean | std dev. | SNR | mean | std dev. | SNR |
| 100 | 103.76 | 18.37 | 5.65 | 33.68 | 5.76 | 5.85 |
| 150 | 102.45 | 18.02 | 5.68 | 33.60 | 5.49 | 6.12 |
| 200 | 103.09 | 17.65 | 5.84 | 33.39 | 5.14 | 6.50 |
| 250 | 102.99 | 16.77 | 6.14 | 33.24 | 4.77 | 6.97 |
| 300 | 103.28 | 16.19 | 6.38 | 33.16 | 4.62 | 7.18 |
| 350 | 103.37 | 15.73 | 6.57 | 33.06 | 4.39 | 7.53 |
| 400 | 103.24 | 15.35 | 6.72 | 32.99 | 4.26 | 7.74 |

**Table 4.** Low-resolution reconstruction without registration.

| No. of | Region 1 | | | Region 2 | | | Vol. |
|---|---|---|---|---|---|---|---|
| B-scans | mean | std dev. | SNR | mean | std dev. | SNR | (ml) |
| 100 | 100.39 | 11.07 | 9.07 | 33.22 | 3.78 | 8.79 | 7.20 |
| 150 | 100.09 | 10.63 | 9.42 | 33.15 | 3.41 | 9.72 | 7.17 |
| 200 | 99.94 | 9.77 | 10.23 | 33.07 | 3.07 | 10.77 | 7.19 |
| 250 | 100.05 | 9.27 | 10.79 | 33.00 | 2.88 | 11.46 | 7.27 |
| 300 | 99.93 | 8.67 | 11.53 | 32.97 | 2.86 | 11.53 | 7.31 |
| 350 | 99.71 | 8.16 | 12.22 | 32.93 | 2.82 | 11.68 | 7.43 |
| 400 | 99.44 | 7.81 | 12.73 | 32.90 | 2.80 | 11.75 | 7.56 |

The volume of the phantom, as estimated by semi-automatic segmentation, increases with the number of B-scans. This is because the registration errors blur the boundary of the phantom, affecting the semi-automatic segmentation.

$\mathbf{T}^*$ is used to transform the new B-scan image **P** into the coordinates of **C**. The pixels of the registered image can now be added to **C**. Where the registered scan plane intersects already-filled voxels, weighted averaging is used to compound the new image with the existing data (see step 2.3 in Figure 3). Figure 13 shows slices through a volume reconstructed in this manner.

**Table 5.** Low-resolution reconstruction with registration.

| No. of | Region 1 | | | Region 2 | | | Vol. |
|---|---|---|---|---|---|---|---|
| B-scans | mean | std dev. | SNR | mean | std dev. | SNR | (ml) |
| 100 | 100.39 | 11.07 | 9.07 | 33.22 | 3.78 | 8.79 | 7.20 |
| 150 | 100.09 | 10.63 | 9.42 | 33.15 | 3.41 | 9.72 | 7.16 |
| 200 | 99.94 | 9.77 | 10.23 | 33.07 | 3.07 | 10.77 | 7.19 |
| 250 | 99.91 | 9.09 | 10.99 | 33.00 | 2.84 | 11.62 | 7.20 |
| 300 | 100.02 | 8.13 | 12.30 | 32.96 | 2.78 | 11.86 | 7.14 |
| 350 | 100.06 | 7.76 | 12.89 | 32.91 | 2.68 | 12.28 | 7.20 |
| 400 | 100.17 | 7.27 | 13.78 | 32.86 | 2.57 | 12.79 | 7.18 |

The volume of the phantom, as estimated by semi-automatic segmentation, remains almost constant with higher levels of compounding.

## 4. RESULTS

### 4.1. Phantom study

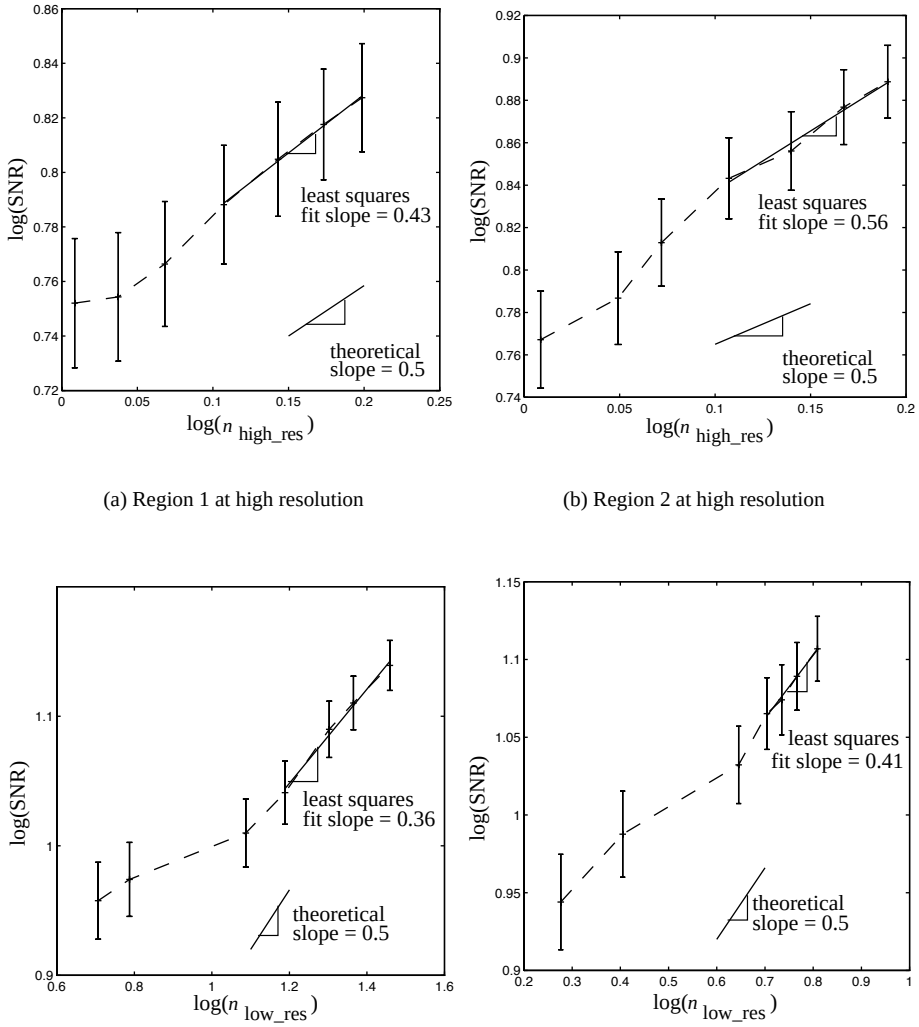
*4.1.1. 3-D spatial compounding and its effect on the SNR*

The main purpose of the phantom study is to investigate the effect of compounding on the SNR of the 3-D reconstructions. The B-scans of the phantom cross-sections contain two regions with almost homogeneous statistical features: the inside of the latex balloon and the outside. Figure 9 shows the two regions on a typical B-scan of the phantom.

The phantom was scanned in a continuous series of sweeps from one end of the balloon to the other, producing a large number of overlapping B-scans. Each sweep was carried out with the probe at a look direction slightly displaced from the previous sweep. About 4–5 sweeps were performed, giving 400 B-scans in total. The first 100 B-scans correspond to the first complete sweep, so they do not overlap each other.

Volumes were reconstructed from the 400 B-scans at both high and low resolution. The high-resolution reconstructions featured cubic voxels with the same edge dimension as 1 pixel in the original B-scan (0.14 mm). The low-resolution reconstructions contained voxels with an edge dimension seven times the size of a pixel (1 mm). While the high-resolution reconstructions preserve the full resolution of the B-scans, they require a large amount of memory (181 Mbytes) and contain many unfilled voxels in regions between B-scans. Low-resolution reconstructions (0.5 Mbytes with 1 mm voxels) can be generated and visualized in significantly less time. For each resolution, two reconstruction volumes were created, one with registration (using the novel reconstruction algorithm in Figure 5), the other without (using the standard reconstruction algorithm in Figure 3). Four volumes were therefore reconstructed in total.

Registration was, in fact, only performed at low resolution, with the results applied to the high-resolution reconstructions.




(a) Region 1 at high resolution

(b) Region 2 at high resolution

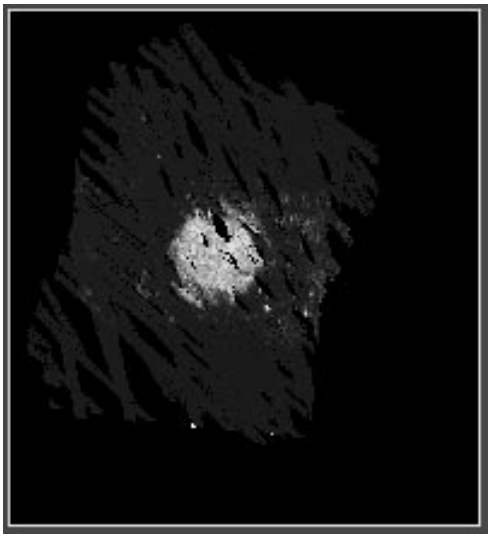
(c) Region 1 at low resolution

(d) Region 2 at low resolution

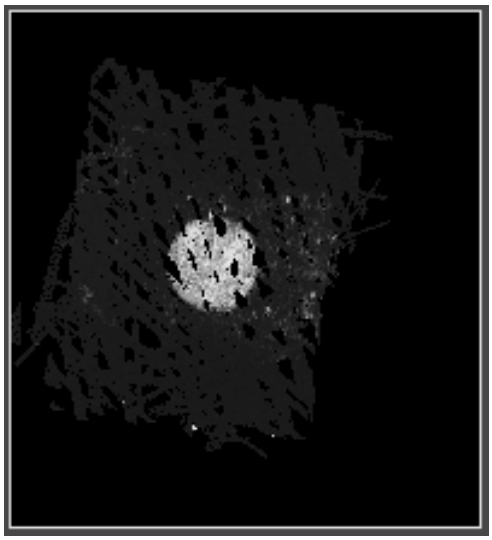
**Figure 10.** Effect of compounding on SNR for registered reconstructions. Both high-resolution (**a** and **b**) and low-resolution (**c** and **d**) reconstructions show an increase in the SNR as the level of compounding increases. Each '+' data point represents an SNR calculated at 100, 150, 200, 250, 300, 350 and 400 B-scans. The error bars were estimated by measuring the RMS variation in the SNR obtained using slightly different volumetric masks. The $x$-axis indices, $n_{\text{high\_res}}$ and $n_{\text{low\_res}}$, are defined in the appendix. A log–log plot is used because a slope of 0.5 indicates agreement with the theoretical increase of SNR as $n^{0.5}$. As explained in Section 4.1.1, only the final four data points in each plot can be meaningfully compared to the theory. In all cases, these data points are in remarkably good agreement with the theory.

Registration at high resolution is infeasible because of the large number of unfilled voxels and the considerable memory requirements of the reconstruction algorithm.

To investigate the improvement in SNR with spatial compounding, we analysed grey-level statistics in each of the four reconstruction volumes at various stages of reconstruction: after 100 B-scans (initial sweep, no compounding), then 150, 200, 250, 300, 350 and finally 400 B-scans (heavy compounding). For each case we segmented volumetric regions 1 and 2, and then calculated the mean and standard deviation (and therefore the SNR) for filled voxels in each region: these are tabulated in Tables 2–5. Segmentation was performed by

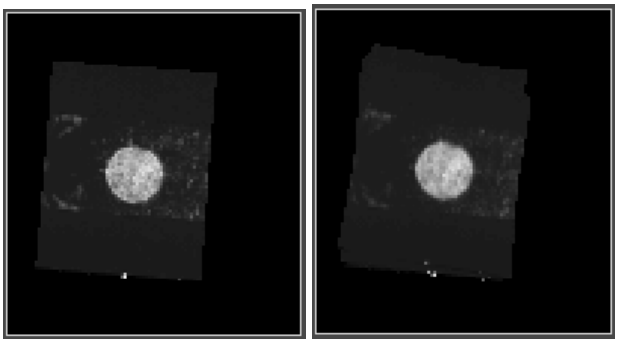


(a) Unregistered

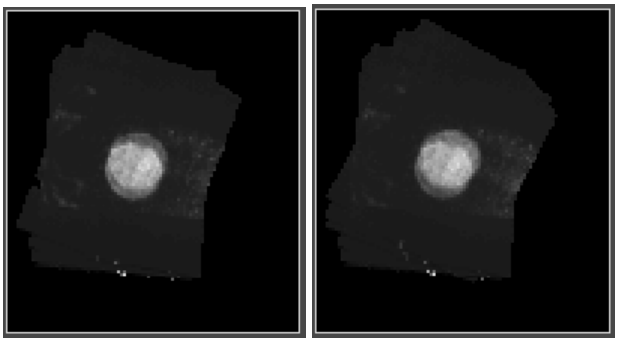


(b) Registered

**Figure 11.** High-resolution reconstruction. A slice of the reconstruction volume is shown for both the unregistered and registered cases. All 400 B-scans are used in both reconstructions, but gaps still remain where voxels are not intersected by any of the B-scans. In (**a**), the registration errors substantially distort the circular cross-section of the phantom. The circular shape is restored in (**b**) by landmark registration.



(a) 100 B-scans used (b) 200 B-scans used



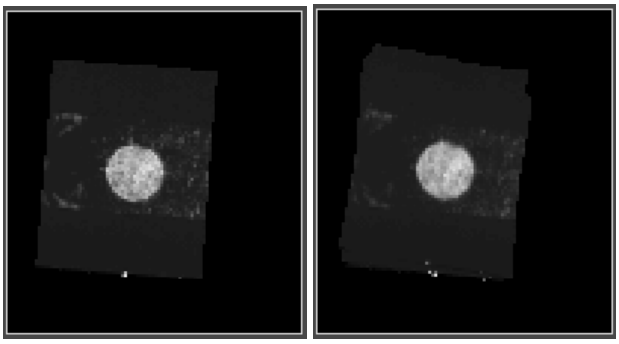
(c) 300 B-scans used (d) 400 B-scans used

**Figure 12.** Low-resolution reconstruction without registration. Figures (**a**–**d**) are of slices taken at the same location in the reconstruction volume at increasing levels of compounding. Speckle is reduced both inside and outside the object, but registration errors result in substantial blurring of the object boundary.

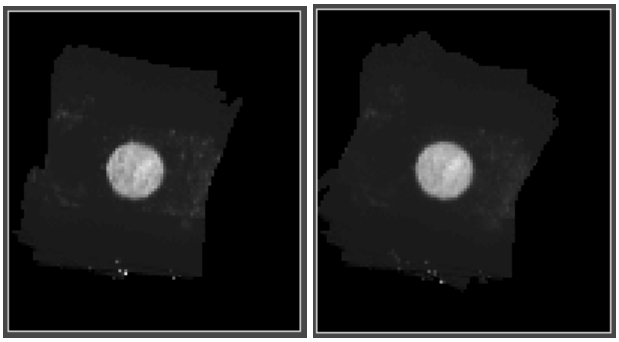
manually constructing volumetric masks for regions 1 and 2. Once each mask was created, it was used for all seven stages of reconstruction (100–400 B-scans). In this way, segmentation differences were not a factor in the observed changes in SNR.

It is immediately apparent that the SNR increases with the amount of compounding. Furthermore, the SNR improves almost identically for both the registered and unregistered cases. This is because regions 1 and 2 do not include the area where the phantom boundary is blurred by the registration errors.

The improvement in SNR can be predicted by statistical theory. Previous papers (Burckhardt, 1978; Kerr *et al.*, 1986; Hernandez *et al.*, 1996) have demonstrated a $\sqrt{n}$ improvement in SNR for 2-D compounding of $n$ uncorrelated

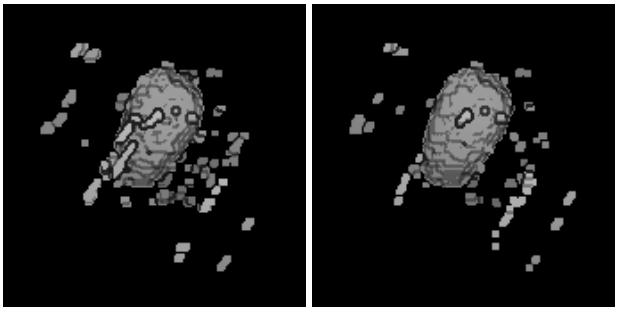


(a) 100 B-scans used (b) 200 B-scans used

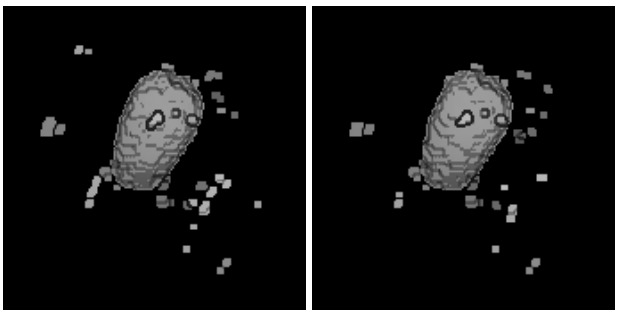


(c) 300 B-scans used (d) 400 B-scans used

**Figure 13.** Low-resolution reconstruction with registration. Figures (**a**–**d**) are of slices taken at the same location in the reconstruction volume at increasing levels of compounding. Speckle is reduced both inside and outside the object and blurring of the object boundary is minimal.



(a) 100 B-scans used (b) 200 B-scans used



(c) 300 B-scans used (d) 400 B-scans used

**Figure 14.** Low-resolution reconstruction with registration. The 3-D surface renderings show how the speckle outside the object is reduced with increasing levels of compounding. (**a**) shows the reconstruction after a single sweep. The object is less obscured by speckle when the reconstruction volume is heavily compounded in (**d**).

B-scans[a]. This theory is not directly applicable to the 3-D case, since the voxels are not all compounded the same number of times. The arbitrary positions and orientations of the B-scans result in some voxels being intersected more than others. Furthermore, the statistical theory for low-resolution compounding must account for both the compounding due to the intersections of multiple B-scans as well as the reduction in resolution. For these reasons we have developed a full theory of 3-D spatial compounding, which can be found in the appendix.

The measured SNR of both the low- and high-resolution registered reconstructions is plotted against the theoretical increase in SNR in Figure 10. As the theory predicts, the SNR increases with an increasing level of compounding. Yet there are a number of differences between the idealized theory and the actual empirical study. The first is that no attempt was made to obtain completely uncorrelated B-scans for different sweeps. It has been shown that the look directions for different sweeps must be greater than 0.4 transducer widths apart to obtain completely uncorrelated speckle patterns (Trahey *et al.*, 1986b). To verify this, we deliberately performed the first few sweeps from similar look directions, saving large variations in the look direction for the last few sweeps. The results in Figure 10 confirm the theory: the SNR improves most dramatically for the last few sweeps, when the speckle patterns are less correlated. For this reason, we have highlighted the last four data points in Figure 10 to compare with the predicted values, which were derived assuming completely uncorrelated speckle (see the appendix).

[a]This assumes that the speckle is uncorrelated across the B-scans. Compounding B-scans from very similar look directions will result in a smaller improvement in the SNR because the speckle will be partly correlated.

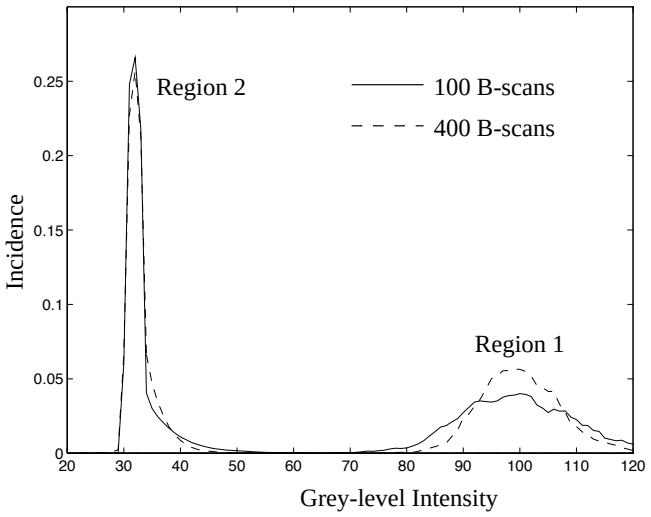



**Figure 15.** Histogram of grey levels in the reconstruction volume with registration. The two regions are described by individual distributions. As compounding increases, the mean grey level remains constant but the standard deviation decreases. This effect is particularly visible for the distribution of region 1.

The grey-level variations in the B-scans are also not due entirely to speckle. The echo amplitudes reflected by a homogeneous medium have an expected Rayleigh distribution when there is a large number of scatterers per resolution cell. The SNR of a Rayleigh distribution is constant and equal to 1.91 (Burckhardt, 1978). The logarithmic compression of echo amplitude in B-scan formation changes the Rayleigh distribution to a Gaussian one, with an SNR of $\sim$7.64 (Thijssen *et al.*, 1988). The measured SNR of a typical B-scan in this study is 5.6 for region 1 and 5.4 for region 2. The lower SNR means that the observed grey-level variations are greater than the variations due only to speckle. If some of the variation is due to real physical structure (such as the fine structure of the talcum powder suspension), then it will be correlated across different sweeps and cannot be reduced by compounding. Despite these differences from the assumptions underpinning the statistical theory, the slopes of the curves in Figure 10 are in remarkably close agreement with the theoretical predictions.

#### *4.1.2. Registration errors and their effect on volume estimation*

Cross-sectional slices of the high-resolution reconstructions are shown in Figure 11, both with and without registration. It is evident that the registration errors are large enough to significantly distort the reconstruction, but are dramatically reduced by landmark-based registration. The effect of compounding is not pronounced at high resolution, because each voxel is intersected only a small number of times.

Figures 12 and 13 show the effects of registration and compounding at low resolution. In both figures the speckle (and other artefacts) are greatly reduced by compounding. However, the shape of the phantom in the unregistered case departs significantly from the original shape. The registered reconstruction maintains the original shape. Figure 14 shows how compounding improves a surface rendering of the reconstruction volume.

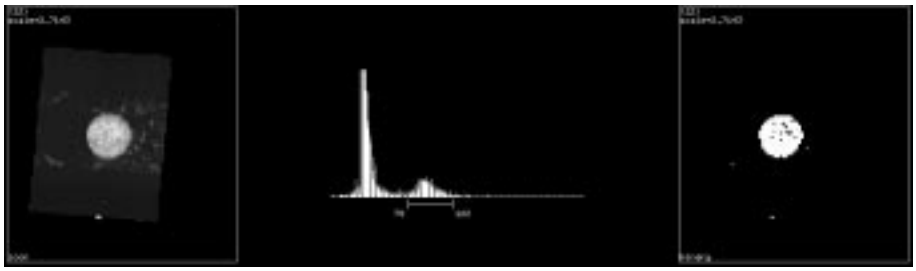
A measure is required to quantify the effect of registration errors on the reconstruction. We chose to focus on the volume of the phantom, as estimated by 'live-wire' segmentation (Barrett and Mortensen, 1996) of slices through the reconstructed volume[a]. This is also a measure that is often sought after by physicians when scanning internal organs. Furthermore, it can be compared with the real volume of the phantom, which was measured with a graduated cylinder at 7.0 ml $\pm$ 0.2 ml.

The volume of the phantom was calculated by live-wire segmentation at seven levels of compounding: after 100, 150, 200, 250, 300, 350 and 400 B-scans. The resulting volume estimates are given in Tables 4 and 5 for the low-resolution reconstructions. Live-wire segmentation is not feasible at high resolution, where there are too many gaps around the boundary.

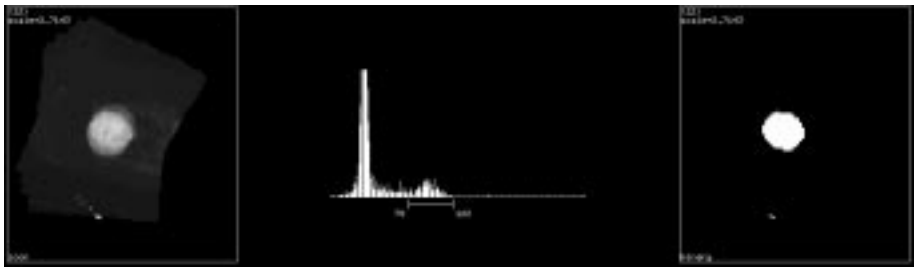
Without registration, the blurring of the phantom boundary results in an increase in the segmented volume (range = [7.17–7.56]) that does not reflect the true volume. Conversely, the volume calculated from the registered reconstructions is stable (range = [7.14–7.20]). The increase in volume by blurring is less dramatic than would be expected from looking at slices through the reconstruction (Figure 12). This is because the live wire is attracted to the location with the highest intensity gradient, which is sometimes near the true boundary and not the edge of the blurred region. It is fair to say that considerably worse volume estimates can be expected under less favourable circumstances.

The grey-level histograms of regions 1 and 2, shown in Figure 15, change as the level of compounding increases. As expected, the histograms become narrower with higher levels of compounding. If the histograms of the different regions are non-overlapping, then fully automatic segmentation can be achieved by thresholding the reconstruction at the appropriate grey level. Figure 16 shows several examples of fully automatic segmentation by grey-level thresholding. One of

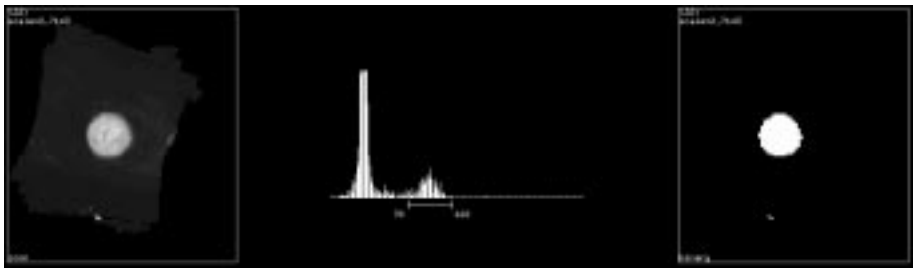
[a]Live-wire segmentation is a powerful tool for extracting boundaries in noisy images. It offers a good compromise between accuracy and amount of user intervention. The technique involves laying an active wire around the object (on a slice-by-slice basis) that is attracted automatically to the object's boundary. The operator assists the live wire by depositing small sections at a time near the boundary, so that the wire does not enclose nearby speckle. For these tests, four sections of wire were sufficient to enclose the phantom cross-section accurately.



(a) Reconstruction using 100 B-scans. The area of the thresholded image on the right is 243.9 $mm^2$. Live-wire semi-automatic segmentation of the image produces an area estimate of 256.5 $mm^2$. The difference arises mainly from speckle in the interior that falls outside the threshold range.



(b) Unregistered reconstruction using 400 B-scans. The area of the thresholded image on the right is 217.7 $mm^2$. The area is lower than in (a) because the blurred regions near the phantom boundary fall outside the threshold range.



(c) Registered reconstruction using 400 B-scans. The area of the thresholded image on the right is 250.9 $mm^2$. This is close to the area calculated by live-wire segmentation in (a).

**Figure 16.** Low-resolution reconstruction. Corresponding slices through the reconstruction volumes are shown on the left of each figure. For each slice, the histogram of the grey levels is shown in the middle. The images on the right show the regions segmented by thresholding the grey levels falling into the range [78–122]. (a) The grey-level segmentation in (**c**) is more accurate than in (**a**) or (**b**). The compounding in (c) ensures that the entire object falls within the threshold range and the registration retains the original circular shape of the cross-section.

the motivations for improving SNR is to allow more accurate automatic segmentation. It is evident that compounding with registration improves the accuracy of automatic grey-level segmentation.

### 4.2. *In vivo* study

To demonstrate that the registered reconstruction algorithm (Figure 5) can be applied to *in vivo* images, we performed an ultrasound examination of the gall bladder of a healthy human subject. Reconstructions were performed at low resolution.

Figure 17 shows that *in vivo* registration errors are significant but can be minimized by landmark-based registration. A more subtle effect is that the high-intensity cloud-like artefact is slightly reduced by compounding. The low acquisition rate of the video framegrabber restricted this study to only 60 B-scans. With higher numbers of compounded B-scans, the artefacts should be further reduced, as in the phantom study.

## 5. CONCLUSIONS

We have shown how spatial compounding can improve the SNR of 3-D ultrasound reconstructions. The improvement in SNR agrees with the theoretical predictions for both the high- and low-resolution cases. The inevitable registration errors which come with higher levels of compounding can be corrected using an automatic, incremental landmark-based registration algorithm. The resulting high quality 3-D reconstructions are particularly well suited to automatic segmentation for visualization and volume measurement.

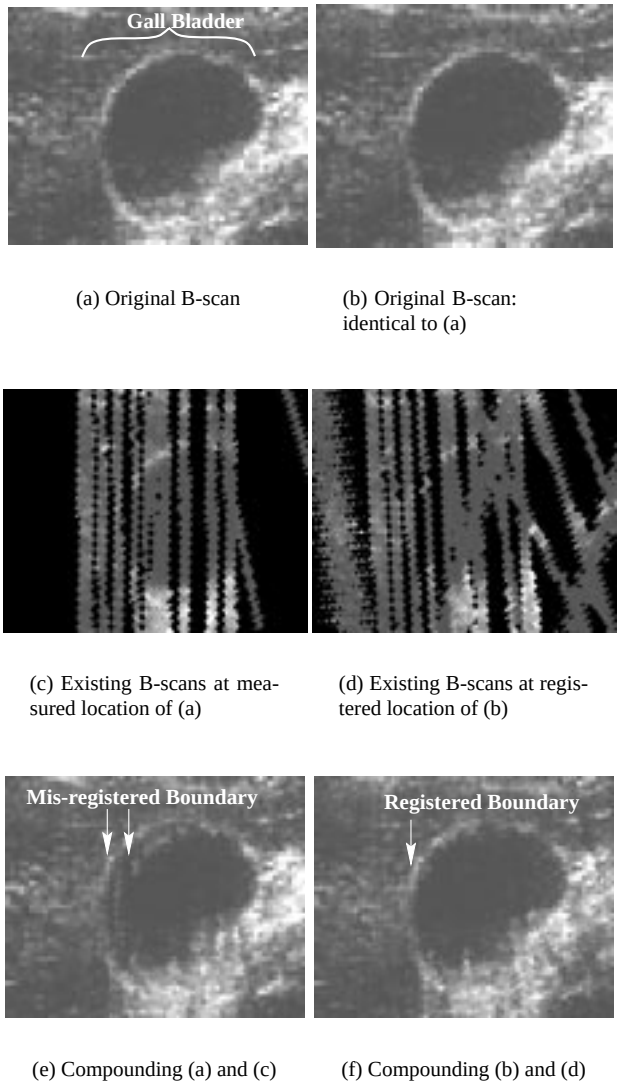
## 6. FUTURE WORK

At present, the reconstruction algorithm only performs registration with rigid-body transformations. We have found that the majority of registration errors can be corrected this way. While the current reconstruction algorithm is sufficiently robust to tolerate small non-rigid errors, these errors remain in the final compounded reconstruction. If larger non-rigid errors occur (for example, from elastic organ movement due to breathing) a non-rigid transformation will be required for accurate registration. We intend to investigate non-rigid registration in future *in vivo* work.

Future work also will investigate higher levels of compounding in *in vivo* scans. We have recently installed a new data acquisition system which can acquire more than 1000 B-scans at 25 frames per second in a single examination. It is likely that the speed and robustness of the registration algorithm will have to be improved to handle the higher levels of compounding. We anticipate that the key to efficient registration will lie with more reliable landmark detection at larger scales, allowing registration of contours in the B-scans onto surfaces in the voxel array. Results of future work will be made available on our WWW site: http://svr-www.eng.cam.ac.uk/~rnr20/research.html

## ACKNOWLEDGEMENTS

The 3-D renderings were produced using 3DViewnix. Robert Rohling is supported by Churchill College and an ORS award.




(a) Original B-scan

(b) Original B-scan: identical to (a)

(c) Existing B-scans at measured location of (a)

(d) Existing B-scans at registered location of (b)

(e) Compounding (a) and (c)

(f) Compounding (b) and (d)

**Figure 17.** Registration of an *in vivo* transverse and longitudinal gall bladder examination. The left-hand column illustrates compounding without registration, the right-hand column with registration. The right-hand column shows the elimination of a double boundary, evident in the left-hand column, that is produced by misregistered B-scans. The high-intensity cloud-like artefact to the right and below the gall bladder is slightly suppressed by compounding. The effect is small because only a few B-scans are compounded in that region. The level of compounding is not high: 35% of filled voxels are intersected by one B-scan, 33% are intersected by two, 16% by three and 16% by four or more. This figure is produced with 1 $mm^3$ voxels to enhance clarity, but registration was performed with $2 \times 2 \times 2$ mm voxels.

## REFERENCES


Barrett, W. A. and Mortensen, E. N. (1996) Fast, accurate, and reproducible live-wire boundary extraction. In Hohne, K. H. and Kikinis, R. (eds), *Lecture Notes in Computer Science: Visualization in Biomedical Computing, Proc. VBC '96*, Vol. 1131, pp. 183–192. Springer-Verlag, Heidelberg.

Brown, L. G. (1992) A survey of image registration techniques. *ACM Comput. Surveys*, 24, 325–376.

Burckhardt, C. B. (1978) Speckle in ultrasound B-mode scans. *IEEE Trans. Sonics Ultrasonics*, 25, 1–6.

Canny, J. (1986) A computational approach to edge detection. *IEEE Trans. PAMI*, 8, 679–698.

Chervenak, F. A., Isaacson, G. C. and Campbell, S. (1993) *Ultrasound in Obstetrics and Gynecology*. Little, Brown and Company, Boston, MA.

Crawford, D. C., Bell, D. S. and Bamber, J. C. (1993) Compensation for the signal processing characteristics of ultrasound B-mode scanners in adaptive speckle reduction. *Ultrasound Med. Biol.*, 19, 469–485.

Downey, D. B., Nicolle, D. A., Levin, M. F. and Fenster, A. (1996) Three-dimensional ultrasound imaging of the eye. *Eye*, 10, 75–81.

Evans, A. N. and Nixon, M. S. (1995) Mode filtering to reduce ultrasound speckle for feature extraction. *IEE Proc. Vision, Image Signal Processing*, 142, 87–94.

Fine, D., Perring, S., Herbetko, J., Hacking, C. N., Fleming, J. S. and Dewbury, K. C. (1991) Three-dimensional (3D) ultrasound imaging of the gallbladder and dilated biliary tree: reconstruction from real-time B-scans. *Brit. J. Radiology*, 64, 1056–1057.

Fischler, M. A. and Bolles, R. C. (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24, 381–395.

Franseschi, D., Bondi, J. A. and Rubin, J. R. (1992) A new approach for three-dimensional reconstruction of arterial ultrasonography. *J. Vasc. Surg.*, 15, 800–805.

Gilja, O. H., Smievoll, A. I., Thune, N., Matre, K., Hausken, T., Odegaard, S. and Berstad, A. (1995) *In vivo* comparison of 3D ultrasonography and magnetic resonance imaging in volume estimation of human kidneys. *Ultrasound Med. Biol.*, 21, 25–32.

Hernandez, A., Basset, O., Chirossel, P. and Gimenez, G. (1996) Spatial compounding in ultrasonic imaging using an articulated scan arm. *Ultrasound Med. Biol.*, 22, 229–238.

Hogg, R. V. and Ledolter, J. (1987) *Engineering Statistics*. Macmillan Publishing Company, New York.

Hottier, F. and Collet Billon, A. (1990) 3D echography: status and perspective. In Hohne, K. H., Fuchs, H. and Pizer, S. M. (eds), *3D Imaging in Medicine: Algorithms, Systems, Applications*, pp. 21–41. Springer-Verlag, Berlin.

Hughes, S. W., D'Arcy, T. J., Maxwell, D. J., Chiu, W., Milner, A., Saunders, J. E. and Sheppard, R. J. (1996) Volume estimation from multiplanar 2D ultrasound images using a remote electromagnetic position and orientation sensor. *Ultrasound Med. Biol.*, 22, 561–572.

Kerr, A. T., Patterson, M. S., Foster, F. S. and Hunt, J. W. (1986) Speckle reduction in pulse echo imaging using phase insensitive and phase sensitive signal processing techniques. *Ultrasonic Imag.*, 8, 11–28.

Leeman, S. and Seggie, D. A. (1987) Speckle reduction via phase. In Ferrari, L. A. (ed.), *Int. Symp. on Pattern Recognition and Acoustical Imaging*, pp. 173–177. SPIE Vol. 768.

Moskalik, A., Carson, P. L., Meyer, C. R., Fowlkes, J. B., Rubin, J. M. and Roubidoux, M. A. (1995) Registration of three-dimensional compound ultrasound scans of the breast for refraction and motion correction. *Ultrasound Med. Biol.*, 21, 769–778.

Nelson, T. R. and Pretorius, D. H. (1994) 3D ultrasound image quality improvement using spatial compounding and 3D filtering. *Med. Phys.*, 21, 998.

Rankin, R. N., Fenster, A., Downey, D. B., Munk, P. L., Levin, M. F. and Vellet, A. D. (1993) Three-dimensional sonographic reconstruction: techniques and diagnostic applications. *Am. J. Roentgenology*, 161, 695–702.

Rohling, R. N. and Gee, A. H. (1996) Issues in 3-D free-hand medical ultrasound imaging. *Technical Report* CUED/F-INFENG/TR 246, Cambridge University Department of Engineering.

Sakas, G., Schreyer, L.-A. and Grimm, M. (1995) Preprocessing and volume rendering of 3D ultrasonic data. *IEEE Comp. Graphics Appl.*, 15, 47–54.

Salustri, A. and Roelandt, J. R. T. C. (1995) Ultrasonic three-dimensional reconstruction of the heart. *Ultrasound Med. Biol.*, 21, 281–293.

Shankar, P. M. (1986) Speckle reduction in ultrasound B-scans using weighted averaging in spatial compounding. *IEEE Trans. UFFC*, 33, 754–758.

Smith, S. W., Davidsen, R. E., Emery, C. D., Goldberg, R. L. and Light, E. D. (1995) Update on 2-D array transducers for medical ultrasound. In *Proc. IEEE Ultrasonics Symp., 1995*, pp. 1273–1278.

Steiner, H., Staudach, A., Spitzer, D. and Schaffer, H. (1994) Three-dimensional ultrasound in obstetrics and gynaecology: technique, possibilities and limitations. *Human Reproduction*, 9, 1773–1778.

Thijssen, J. M., Oosterveld, B. J. and Wagner, R. F. (1988) Gray level transforms and lesion detectability in echographic images. *Ultrason. Imag.*, 10, 171–195.

Trahey, G. E., Allison, J. W., Smith, S. W. and von Ramm, O. T. (1986a) A quantitative approach to speckle reduction via frequency compounding. *Ultrasonic Imag.*, 8, 151–164.

Trahey, G. E., Smith, S. W. and von Ramm, O. T. (1986b) Speckle pattern correlation with lateral aperture translation: experimental results and implications for spatial compounding. *IEEE Trans. UFFC*, 33, 257–264.

Trobaugh, J. W., Trobaugh, D. J. and Richard, W. D. (1994) Three-dimensional imaging with stereotactic ultrasonography. *Computerized Med. Imag. Graphics*, 18, 315–323.

## APPENDIX A. STATISTICAL THEORY OF 3-D SPATIAL COMPOUNDING

Simple statistical theory can be used to predict the increase in SNR from 3-D spatial compounding. In this appendix we consider two distinct cases: 3-D high-resolution compounding and 3-D low-resolution compounding. First, however, we review some simple results from statistical theory.

(i) **Linear functions of random variables** (Hogg and Ledolter, 1987). Consider the random variable $Y$ as a linear function of independent random variables $X_i$ : $Y = \sum_{i=1}^{n} a_i X_i$ where $\{a_1 \dots a_n\}$ are constant coefficients. Then the expectation (or mean) $\mu_Y$ of $Y$, expressed in terms of the expectation of $X_i (\mu_{X_i})$, is $\mu_Y = \sum_{i=1}^{n} a_i \mu_{X_i}$. The variance of $Y$ $(\sigma_Y^2)$, in terms of the variance of $X_i$ $(\sigma_{X_i}^2)$, is $\sigma_Y^2 = \sum_{i=1}^{n} a_i^2 \sigma_{X_i}^2$.

(ii) **Sets of samples of random variables.** Consider $Z$ as a set of samples $Y_j$ drawn randomly from a number of independent distributions: $Z = \{Y_1, \dots, Y_j, \dots, Y_m\}$. The expected mean of the samples in $Z$, expressed in terms of the expectation $(\mu_{Y_j})$ of the distribution from which each $Y_j$ is drawn, is $\mu_Z = \frac{1}{m} \sum_{j=1}^{m} \mu_{Y_j}$. Provided all $\mu_{Y_j}$ are equal, the expected variance of the samples in $Z$, expressed in terms of the variances $(\sigma_{Y_j}^2)$ of the distributions from which each $Y_j$ is drawn, is $\sigma_Z^2 = \frac{1}{m} \sum_{j=1}^{m} \sigma_{Y_j}^2$. These equations are readily derived from the definitions of expectancy and variance.

### A.1. 3-D high-resolution compounding

The simplest 3-D compounding case occurs when the voxel size in the reconstruction is the same as the pixel size of the B-scan. This means that each voxel is intersected no more than once per B-scan (discounting the voxels that can occasionally contain 2 pixels if intersected obliquely). To predict the improvement in SNR, we need to make two key assumptions:

**Assumption 1.** The subject of the ultrasound examination exhibits no resolvable structure, so any grey-level variations in the B-scans are due entirely to speckle.

**Assumption 2.** Separate B-scans are taken from look directions spaced sufficiently far apart, so that the speckle across B-scans is uncorrelated.

Now consider $Z$ to be the 3-D array of voxels $Y_j$, each compounded from pixels $X_i$ of the individual B-scans. Each $X_i$ comes from a distribution with mean $\mu_0$ and variance $\sigma_0^2$. The $X_i$ averaged for a particular $Y_j$ are independent, because the speckle is uncorrelated across B-scans. There are $m$ voxels in the set $Z$ and voxel $j$ is intersected $n_j$ times. The maximum possible value of $n_j$ is the total number of B-scans used in the reconstruction:

$$Z = \{Y_1, \dots, Y_j, \dots, Y_m\} \qquad \text{where} \qquad Y_j = \sum_{i=1}^{n_j} \frac{1}{n_j} X_i \,.$$

From (1),

$$\mu_{Y_j} = \sum_{i=1}^{n_j} \frac{1}{n_j} \mu_0 = \mu_0 \qquad \text{and} \qquad \sigma_{Y_j}^2 = \sum_{i=1}^{n_j} \frac{\sigma_0^2}{n_j^2} = \frac{\sigma_0^2}{n_j}$$

From (2),

$$\mu_Z = \frac{1}{m} \sum_{j=1}^{m} \mu_0 = \mu_0 \qquad \text{and} \qquad \sigma_Z^2 = \frac{1}{m} \sum_{j=1}^{m} \frac{\sigma_0^2}{n_j}.$$

Combining these results gives

$$\frac{\mathrm{SNR}(Z)}{\mathrm{SNR}(Z_0)} = \frac{\mu_Z / \sigma_Z}{\mu_{Z_0} / \sigma_{Z_0}} = \sqrt{\frac{1}{\frac{1}{m} \sum_{j=1}^{m} \frac{1}{n_j}}} \,.$$

To abbreviate references to this result in the main text, let us define the quantity $n_{\text{high\_res}}$ as follows:

$$n_{\text{high\_res}} = \frac{1}{\frac{1}{m} \sum_{j=1}^{m} \frac{1}{n_j}} \qquad \Rightarrow \qquad \frac{\mathrm{SNR}(Z)}{\mathrm{SNR}(Z_0)} = \sqrt{n_{\text{high\_res}}}.$$

Note that the $\sqrt{n}$ result of the 2-D case can be derived by setting $n_j = n$ for all $j$.

### A.2. 3-D low-resolution compounding

Most 3-D reconstructions involve cuberilles with voxels considerably larger than the pixels in the B-scans. Again, the set of voxels in the reconstruction can be written as

$$Z = \{Y_1, \dots, Y_j, \dots, Y_m\} \qquad \text{where} \qquad Y_j = \sum_{i=1}^{n_j} \frac{1}{n_j} X_i$$

but in this case the $X_i$ are not all independent, since some are from the same neighbourhood in the same B-scan (from the reduction in resolution). This becomes clear if we consider each voxel $Y_j$ as follows:

$$Y_j = \sum_{k=1}^{k_j} \sum_{p=1}^{p_{j,k}} \frac{1}{n_j} X_{p,k}$$

where $k_j$ is the number of B-scans that intersect voxel $j$ and $p_{j,k}$ is the number of pixels that intersect voxel $j$ for B-scan $k$: i.e. $\sum_{k=1}^{k_j} p_{j,k} = n_j$. We have labelled the individual pixels $X_{p,k}$ according to B-scan $k$ and pixel $p$. Rearranging gives

$$Y_j = \sum_{k=1}^{k_j} \frac{p_{j,k}}{n_j} \Theta_{k,j} \qquad \text{where} \qquad \Theta_{k,j} = \sum_{p=1}^{p_{j,k}} \frac{1}{p_{j,k}} X_{p,k}$$

and $\Theta_{k,j}$ is the mean of the pixels that intersect voxel $j$ for B-scan $k$. The expected mean ($\mu_\Theta$) of $\Theta_{k,j}$ can be determined using the result for the expectation in (1), which is still valid for non-independent $X_{p,k}$ (Hogg and Ledolter, 1987):

$$\mu_\Theta = \sum_{p=1}^{p_{j,k}} \frac{1}{p_{j,k}} \mu_0 = \mu_0 \,.$$

Since the size of the speckle is larger than 1 pixel, the grey-level value of a pixel is not independent of its close neighbours. Thus the expected variance ($\sigma_\Theta^2$) of $\Theta_{k,j}$ cannot be determined using the result for the variance in (1), which is not valid if the $X_{p,k}$ are not independent. The variance of the sum of several neighbouring pixels does decrease as the number of summed pixels increases, but $\sigma_\Theta^2 \gg \sigma_0^2/p_{j,k}$. $\sigma_\Theta^2$ depends on the spatial structure of the speckle and can be expressed as follows:

$$\sigma_\Theta^2 = \sigma_0^2 b^2(p_{j,k})$$

where $b(p_{j,k})$ is estimated empirically from the B-scans themselves. Returning now to the expression for $Y_j$, the summation over $k$ can be considered as a sum of independent variables, since each element in the sum comes from a different B-scan. Hence

$$\mu_{Y_j} = \sum_{k=1}^{k_j} \frac{p_{j,k}}{n_j} \mu_0 = \mu_0 \text{ and } \sigma_{Y_j}^2 = \sum_{k=1}^{k_j} \left(\frac{p_{j,k}}{n_j}\right)^2 \sigma_0^2 b^2(p_{j,k}).$$

From (2),

$$\mu_Z = \frac{1}{m}\sum_{j=1}^{m} \mu_0 = \mu_0 \text{ and } \sigma_Z^2 = \frac{1}{m}\sum_{j=1}^{m}\sum_{k=1}^{k_j}\left(\frac{p_{j,k}}{n_j}\right)^2 \sigma_0^2 b^2(p_{j,k}).$$

Combining these results gives

$$\frac{\mathrm{SNR}(Z)}{\mathrm{SNR}(Z_0)} = \frac{\mu_Z/\sigma_Z}{\mu_{Z_0}/\sigma_{Z_0}} = \sqrt{\frac{1}{\frac{1}{m}\sum_{j=1}^{m}\sum_{k=1}^{k_j}\left(\frac{p_{j,k}}{n_j}\right)^2 b^2(p_{j,k})}}\,.$$

Again, to abbreviate references to this result in the main text, let us define the quantity $n_{\mathrm{low\_res}}$ as follows:

$$n_{\mathrm{low\_res}} = \frac{1}{\frac{1}{m}\sum_{j=1}^{m}\sum_{k=1}^{k_j}\left(\frac{p_{j,k}}{n_j}\right)^2 b^2(p_{j,k})} \quad \Rightarrow$$

$$\frac{\mathrm{SNR}(Z)}{\mathrm{SNR}(Z_0)} = \sqrt{n_{\mathrm{low\_res}}}\,.$$

Note that the result of the high-resolution case can be derived by setting $p_{j,k} = 1$, $b^2(p_{j,k}) = 1$ (because $\sigma_\Theta^2 = \sigma_0^2$) and $k_j = n_j$.