HYBRID SYSTEMS FOR RECONSTRUCTION OF FREEHAND 3D ULTRASOUND DATA

Article

4 authors, including:

Richard W Prager
University of Cambridge
196 PUBLICATIONS   4,855 CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   Hybrid Three-Dimensional Ultrasound View project

HYBRID SYSTEMS FOR RECONSTRUCTION OF FREEHAND 3D ULTRASOUND DATA

R. J. Housden, G. M. Treece,
A. H. Gee and R. W. Prager

CUED/F-INFENG/TR 574

March 2007

University of Cambridge
Department of Engineering
Trumpington Street
Cambridge CB2 1PZ
United Kingdom

Email: rjh80/gmt11/ahg/rwp @eng.cam.ac.uk

Hybrid systems for reconstruction of freehand 3D ultrasound data

R. James Housden, Graham M. Treece, Andrew H. Gee and Richard W. Prager

University of Cambridge
Department of Engineering
Trumpington Street
Cambridge CB2 1PZ

Abstract

Freehand 3D ultrasound can be acquired without a position sensor by finding the separations of pairs of frames using information in the images themselves. However, small biases in the offset estimates lead to large scale drift in the final reconstruction. In comparison, position sensors have good large scale accuracy, but are inconvenient to use. In this paper, we present methods that combine existing sensorless techniques with limited information from a position sensor, in order to reduce drift errors. We also consider two novel position sensors that are potentially less inconvenient and we assess their accuracy for the purpose of drift correction.

1 Introduction

3D ultrasound [6, 19] is an emerging medical imaging modality with many potential applications [10]. The data can be acquired using special 3D probes, with either a 2D array of crystals or an internal mechanism to sweep the scan head over a 3D volume. Another possibility is the freehand approach, which uses a conventional 2D probe. The clinician manually sweeps the probe over a volume of interest while a sequence of images is recorded. At the same time, a position sensor attached to the probe labels each image frame with its position and orientation, resulting in a 3D volume of data. The 3D data can then be visualised in various ways to extract clinically useful information.

The freehand approach has certain advantages over 3D probes: it allows arbitrary volumes, is relatively low-cost and works with any standard 2D probe. There are also disadvantages: it is slow and therefore not suitable for 4D scanning and the data can be distorted by varying probe contact pressure. 3D probes are currently the focus of most commercial systems. This is perhaps because of the complexity of the freehand approach, particularly the add-on position sensor. Sensors have only a limited operating range and they often need a direct line of sight between the device attached to the probe and the base unit. Also, careful calibration [17] is required each time the sensor is reattached to the probe. Correct reconstruction requires that there is no movement of the scanning subject relative to the sensor origin. Even if the subject doesn't move, small errors in the sensor measurements combined with small movements of the anatomy due to probe pressure result in a jitter error, which reduces the clarity of reslice images through the 3D volume of data.

In previous work [12], we have developed an alternative to the position sensor, which determines the relative position of two image frames using information in the images themselves. In the plane of the images, the offset is found using standard image registration and motion tracking techniques [2, 24]. Out of the image plane, we use speckle decorrelation [3, 14, 23] to estimate the offset at various points over the frame. Given three non-collinear distance estimates, we can then determine the out-of-plane offset in terms of elevational distance, yaw and tilt angles. By concatenating the offsets between pairs of frames, a sequence of frames can be reconstructed.

The speckle decorrelation method works by calculating the correlation between two patches of data, one from each frame. The correlation value decreases as the separation of the two patches increases, and can therefore be converted to a distance estimate using a calibrated decorrelation curve like that shown in Figure 1. However, it is not only elevational separation that affects the correlation value: tissue properties, tissue compression, tissue motion, electrical noise, in-plane
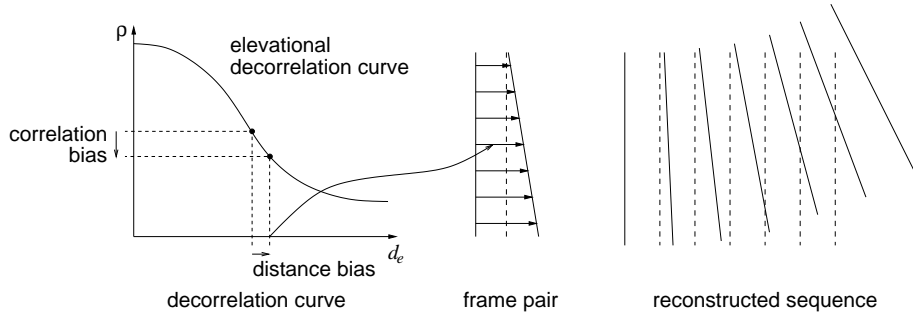
Figure 1: **Reconstruction drift**. The correlation between two patches of data depends on the separation between them. Given a correlation value, this can be converted to a distance. However, various effects can bias the correlation value, which gives biased distance estimates. This results in biased inter-frame offsets and therefore a drift error in the reconstructed sequence. Here, the dashed lines show the correct frame positions and the solid lines show the actual drifted reconstruction.

probe motion and probe rotation can change the correlation. This results in biased distance estimates and therefore a bias in the out-of-plane elevational offset between frame pairs. If the distance bias varies top to bottom over the frame pair, there will also be a bias in the tilt angle. Similarly, there will be a yaw bias if the distance bias varies left to right. Concatenating the individual frame pair offsets through the sequence results in a gradually increasing drift error, as shown in the figure. A consequence of this is that measurements taken from the 3D reconstruction, for example lengths and volumes, will also be biased.

In previous work, we have addressed some of the issues that lead to bias. The basic speckle decorrelation theory only holds when the image consists of fully developed speckle. In [9], we describe an adaptive speckle decorrelation method that adjusts the decorrelation curves to something more suitable for the tissue being scanned. We have also looked at the problem of in-plane probe motion [13]. By correcting for the in-plane motion before estimating the correlation, this source of decorrelation is avoided. We have shown how to improve the accuracy of this correction by interpolating the in-plane shift and correlation value between the sparse lateral A-lines. Both of the above methods have been shown to improve the reconstruction accuracy, but they do not completely remove the drift. In particular, there remains some bias in the elevational separation between frame pairs, as well as the tilt angle.

Clearly, there are advantages and disadvantages to both the sensor systems and the sensorless methods. Sensorless reconstruction is ideal for producing clear qualitative images, because it is not affected by jitter. However, the drift affects the accuracy of measurements taken from the data. Provided the subject is not free to move and the sensor is used in an appropriate environment, position sensors provide accurate large scale measurements that are not affected by drift. In previous work [22], we introduced the idea of using a position sensor to correct for drift when doing image-based in-plane registration. We use the image-based methods to get improved small scale reconstruction accuracy, while the position sensor corrects the large scale drift error. We therefore have a hybrid system, with the best features of each method. In this paper, we present an extension to our sensorless reconstruction algorithms that uses the hybrid approach to correct both the in-plane and out-of-plane drift.

With the reconstruction requiring less information from the position sensor, we are also able to consider some alternative sensor systems. While these are not necessarily as effective as some of the standard sensors, they may be sufficient in a hybrid system and could potentially be incorporated into a freehand 3D ultrasound system with less inconvenience to the user. In this paper, we consider two such sensors. The first is a Measurand ShapeTape[1] [4], which uses optical fibre sensors to
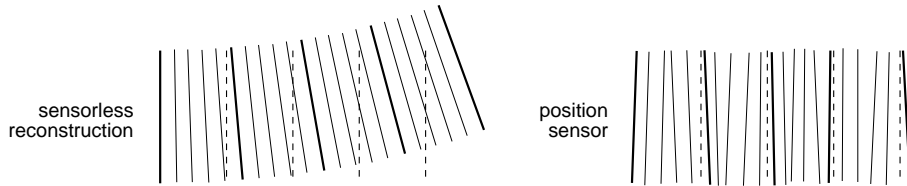
---

[1] http://www.measurand.com

Figure 2: **Drift correction**. The figure shows the frame position data available to the drift correction algorithm. The solid lines are the reconstructed frame positions. The dashed lines show the correct positions. The sensorless reconstruction is done using coarsely spaced frames, shown by the bold lines. The left sequence of frames is the sensorless reconstruction, which has an obvious drift error. The right sequence shows a reconstruction using position sensor measurements.

determine its own shape and therefore its end-point position. The second is an Xsens MT9-B[2], which uses magnetometers, accelerometers and rate gyros to track its orientation. Obviously, this sensor could only be useful in a hybrid system, as it does not provide the complete six degree of freedom (DOF) information.

The paper is structured as follows. In Section 2, we describe our method for correcting drift using information from a position sensor. In Section 3, we describe the two position sensors and our methods for calibrating them to the ultrasound probe. We then assess their accuracy for the purpose of drift correction in a hybrid system. Finally, we conclude in Section 4.

2 Drift correction using position sensors

2.1 Six degree of freedom position sensors

We have explained in the introduction how the drift error in sensorless reconstructions is a consequence of small biases in the individual distance estimates. These are summed up in the reconstruction and the positions gradually drift away from the correct positions, as shown in Figure 2. In comparison, the positions measured by a position sensor do not have a large scale drift error, but have small scale jitter. The objective of the drift correction is to use the large scale sensor information as a reference to correct the sensorless drift, but without the jitter affecting the improved small scale accuracy given by the image-based methods. The correction works by applying a small, constrained adjustment to each of the frame positions relative to the previous frame in the sequence. In the example in the figure, the appropriate adjustment would move each frame so that it is slightly closer and at less of an angle to the previous frame. The combined effect of these small changes significantly reduces the drift error.

The algorithm for correcting drift is designed as an extension to our previous algorithm for sensorless reconstruction [12]. It is therefore necessary to give a brief description of certain features of that algorithm before describing drift correction. First, the reconstruction is done using coarsely spaced frames selected from the sequence, rather than every adjacent frame pair, as shown in Figure 2. This ensures that the distance estimates are taken from the steepest part of the decorrelation curve (see Figure 1) and are therefore more reliable than if closely spaced frames were used [11]. The rest of the frames are added later, without changing the coarse sequence. Secondly, there is a directional ambiguity in the reconstruction, as the speckle decorrelation method only gives the distance and not the direction of the offset. The elevational motion could change direction and this would not show up in the speckle decorrelation distance estimates. By considering three frames at a time, the algorithm is able to overcome this and detect non-monotonic elevational probe motion [12]. Care is taken at the coarse scale to ensure that the sequence is reconstructed with all the offset directions correct. The drift correction described here will therefore only consider the frames in the more reliable coarse sequence when adjusting the reconstruction.

---

[2]http://www.xsens.com

Having done the coarse scale sensorless reconstruction, the position of each frame f in the coarse sequence is expressed as a homogeneous transformation matrix, $\mathbf{S}_f$. The drift correction adjustment is made by premultiplying this by $\mathbf{T}_f$ for each frame.

$$
\begin{aligned}
\mathbf{T}_f &= \mathbf{T}_{f-1}\mathbf{S}_f\mathbf{C}_f\mathbf{S}_f^{-1} \qquad &\text{if} \quad f \geq 1 \\
\mathbf{T}_f &= \mathbf{S}_f\mathbf{C}_0\mathbf{S}_f^{-1} \qquad &\text{if} \quad f = 0
\end{aligned}
\tag{1}
$$

where $\mathbf{C}$ is the adjustment of the frame position relative to the previous frame in the sequence. Note that the adjustment for each frame is applied according to the direction of the bias for that frame pair, i.e. in image frame coordinates. Multiplying by $\mathbf{T}_{f-1}$ takes account of all the adjustments to earlier frames in the sequence.

The question now is how do we decide on $\mathbf{C}$ for each frame. The individual elevational distance estimates tend to be overestimated. We could most likely improve on the elevational bias by applying an adjustment that simply scales the offset between each adjacent frame pair. However, when it comes to the tilt bias, the situation is more complicated. Tilt bias is caused by varying biases in the distance estimates at different heights in the image. Depending on how these biases vary over the height, the tilt angle could be overestimated or underestimated. It is expected that as the probe motion varies through the sequence, the bias between frame pairs will vary. Rather than trying to understand the details of tilt bias, we have chosen to apply a more complex adjustment to the inter-frame offsets than a simple scaling. This will allow for the bias variation through the sequence and should be able to give a better match to the position sensor measurements.

The appropriate transformation matrix $\mathbf{C}$ for each frame is determined using a set of six polynomial curves (two are shown here):

$$
\begin{aligned}
x &= \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \ldots + \alpha_n t^n \\
y &= \beta_0 + \beta_1 t + \beta_2 t^2 + \ldots + \beta_n t^n \\
&\vdots
\end{aligned}
\tag{2}
$$

where t is a value between 0 and 1 defining how far through the sequence the frame is. The value of n determines the degree of flexibility in the warping. For our experiments on drift correction, we set n to 3. According to the value of t, the six equations define the six parameters of the position adjustment (three translations and three rotations) relative to the previous frame in the sequence. We also allow an initial offset at the first frame, defined by an extra six parameters, independent of the correction curves in Equation 2. The overall effect can be thought of as a rigid registration of the sensorless sequence to the correct sequence combined with a non-rigid deformation.

In non-monotonic data, the out-of-plane offset between a frame pair could be reversed during the reconstruction and this will reverse the bias direction. For these frame pairs, we multiply the output from Equation 2 by -1 for the out-of-plane parameters, so that the bias correction is applied in the correct direction.

The coefficients of Equation 2 and the six initial offset parameters are determined by nonlinear optimisation (Levenberg-Marquardt [18]) so that the adjusted sensorless reconstruction is as close as possible to the position sensor measurements, within the limits of the constrained deformation. For each frame, the optimisation error function is defined as the absolute distance between the drift-corrected sensorless positions and the measured positions, at each of the four corners of the frame. This means that errors in the frame orientation as well as the position are detected and minimised.

One final detail is how the value of t is determined for each frame. The root mean square value of the elevational offset at the four corners of a frame pair defines an absolute "distance" between the two frames. These are summed up through the sequence and normalised to the end value. The result is a value that increases monotonically between 0 and 1 through the sequence and this gives t for each frame.
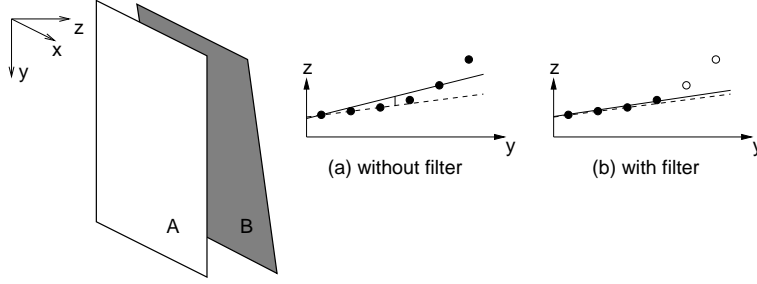
Figure 3: **Filtering incorrect distance estimates**. Biases in the individual distance estimates can lead to biases in the elevational offset and tilt angle parameters when a plane is fitted to the distances, as shown in (a). The dashed line shows the correct plane and the solid line is the plane defined by the elevational distance estimates. By removing distance estimates that are not consistent with the measured orientation, the bias can be reduced, as shown in (b).

2.2 Three degree of freedom orientation sensors

As we mentioned in the introduction, there are sensors available that measure only the orientation and not the position. It is therefore worth considering how such a system could improve on the drift error. We obviously do not have enough information to do a full six DOF correction in the way described above, but we could do something similar to the three orientation parameters. Instead of the six correction curves in Equation 2 and the six initial offset parameters, we have only three of each. The error function for frame orientations is then the in-plane roll, and out-of-plane yaw and tilt angles between the drift-corrected sensorless frame orientation and the measured orientation.

Unlike the six DOF correction, the deformation does not define an entirely new set of frame locations, it only gives a new set of orientations. We need to somehow rotate the frames to the new orientations without affecting the translations established by the sensorless reconstruction. We therefore define $\mathbf{C}$ in Equation 1 to rotate the frame so that the frame centre does not move relative to the previous frame. In our example in Figure 2, the adjustment would decrease the angle between each frame pair, and this alone would move the end frame significantly closer to the correct position.

With this correction, there will still be a bias in measurements taken along the sweep, as there is still an elevational offset bias between frame pairs. However, there are other more subtle ways that the orientation measurements could be used to improve the initial sensorless offset estimates. We use the orientation data as a filter to remove distance estimates with large errors. Figure 3 shows a typical frame pair and the resulting variation in distance estimates down the frame, as determined by speckle decorrelation. In this example, the bias in the distance estimates is larger towards the bottom of the frame. This could be caused by a lack of signal in this region of the image, resulting in lower correlations and hence overestimated distance estimates. Fitting a plane to these distances directly will result in a tilt bias and a small overestimate in the elevational centre offset parameter, as shown in Figure 3(a).

The filter we propose is based on a RANSAC algorithm [7], which will determine which of the estimates should be ignored. It works by considering each estimate individually and fitting a plane to it using the known orientation, measured using the sensor. We then count how many of the other estimates support the hypothesis that this one is correct, determined by how close they are to the fitted plane. The plane that results in the largest support set provides the set of "correct" distance estimates and any others are then ignored. In our example, the distance estimates highlighted in Figure 3(b) would be ignored, resulting in the much better fit shown in the figure. The filter has one tunable parameter: the threshold specifying how close an estimate must be to the fitted plane to count as a supporter. We set the threshold to one standard deviation of the estimated uncertainty in the distances, determined using the method described in [12].
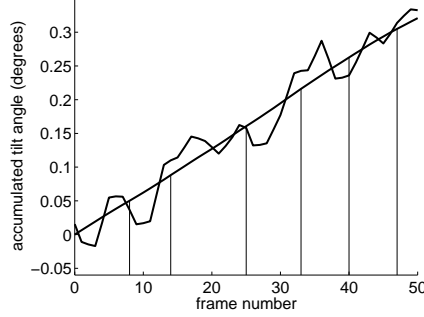
Figure 4: **Smoothed sensor data**. The effect of jitter is obvious when looking at the reconstructed sequence in terms of the accumulated tilt angle through the sequence. The graph shows an example of a jittery set of sensor data, and a smoothed version. The vertical lines show the frames used in the coarse sequence. The jitter affects the measured angle between adjacent frames in the coarse sequence.

Finally, we need to consider the effect of jitter on the RANSAC filtering. Figure 4 shows an example of a jittery reconstruction. The jitter averages out over the length of the sequence, but it is significant between the individual frame pairs used in the RANSAC filter. We reduce the effect of jitter by smoothing the sensor data using a low pass filter.

2.3 Experiments and results

In this section, we present results showing the effectiveness of the drift correction on some *in vitro* data sets. We recorded ten data sets by freehand scanning a beef joint in monotonic and non-monotonic sweeps. We also recorded one additional scan of a speckle phantom, which was intended to be more of a stress test for the reconstruction algorithms. This data set has deliberate rotation as well as non-monotonic elevational motion. For all data sets, the position and orientation of each frame were recorded using a Northern Digital Polaris optical tracking system[3] with a Traxtal Technologies AdapTrax active tool[4] attached to the probe.

The sequences were reconstructed without the sensor information using the methods in [12]. We also repeated the reconstruction using the RANSAC filter and warping enhancements described above, using both the full sensor data and only the orientation data. Figures 5 and 6 show the two drift correction results for each data set. The results are compared to the purely sensorless and the position sensor reconstructions. The results are in terms of the accumulated elevational centre offset between frame pairs through the data set, referred to as the "length" of the sequence, and the accumulated tilt angle between frame pairs.

There are several observations to make from these results. First, the six DOF position sensor does a good job of making the drift-corrected length match the measured length. In all eleven cases, the drift is reduced. The tilt correction is also very good with errors of no more than half a degree anywhere.

It should be noted that what we have achieved here is to get a close match to the position sensor. Whether the sensor measurements are correct or not is another matter. When recording a sweep, care must be taken that the subject does not move relative to the sensor origin. If there is movement, the sensor based reconstruction will be incorrect and cannot be used as a reference for drift correction.

The results for orientation sensor data only do not show such significant improvements. In most cases, there is no noticeable improvement in the length drift. The reason for this must be that the variation in the elevational distance estimate bias is evenly distributed about the centre

---

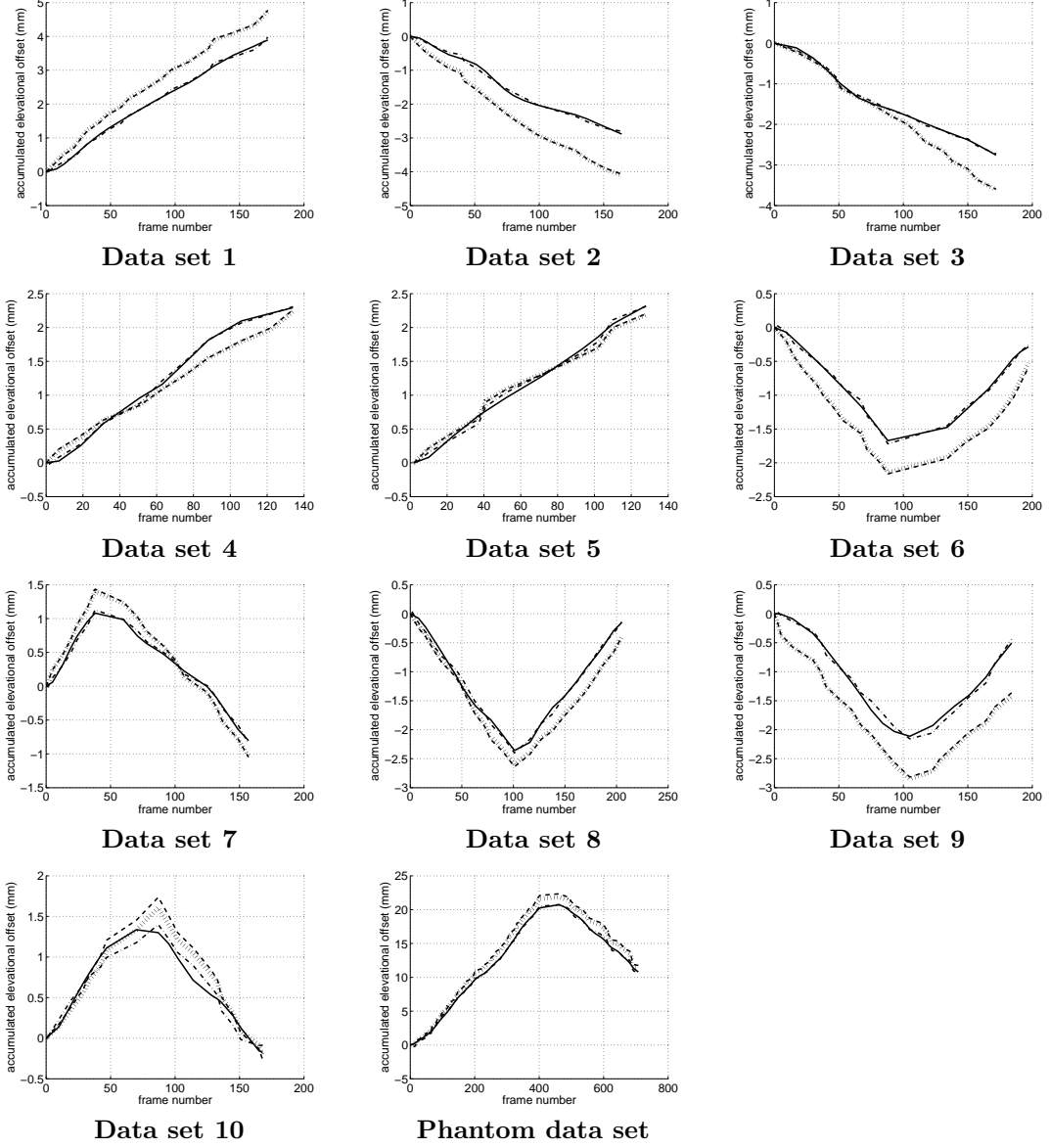[3]http://www.ndigital.com
[4]http://www.traxtal.com

Figure 5: **Drift correction of length**. The graphs show the elevational offset at the centre of the frame accumulated through the data set. The solid lines (—) show smoothed Polaris sensor data. The dashed lines (- -) show the sensorless reconstruction. The six DOF drift-corrected sensorless reconstruction (-··-) and orientation drift correction (||||) are also shown.
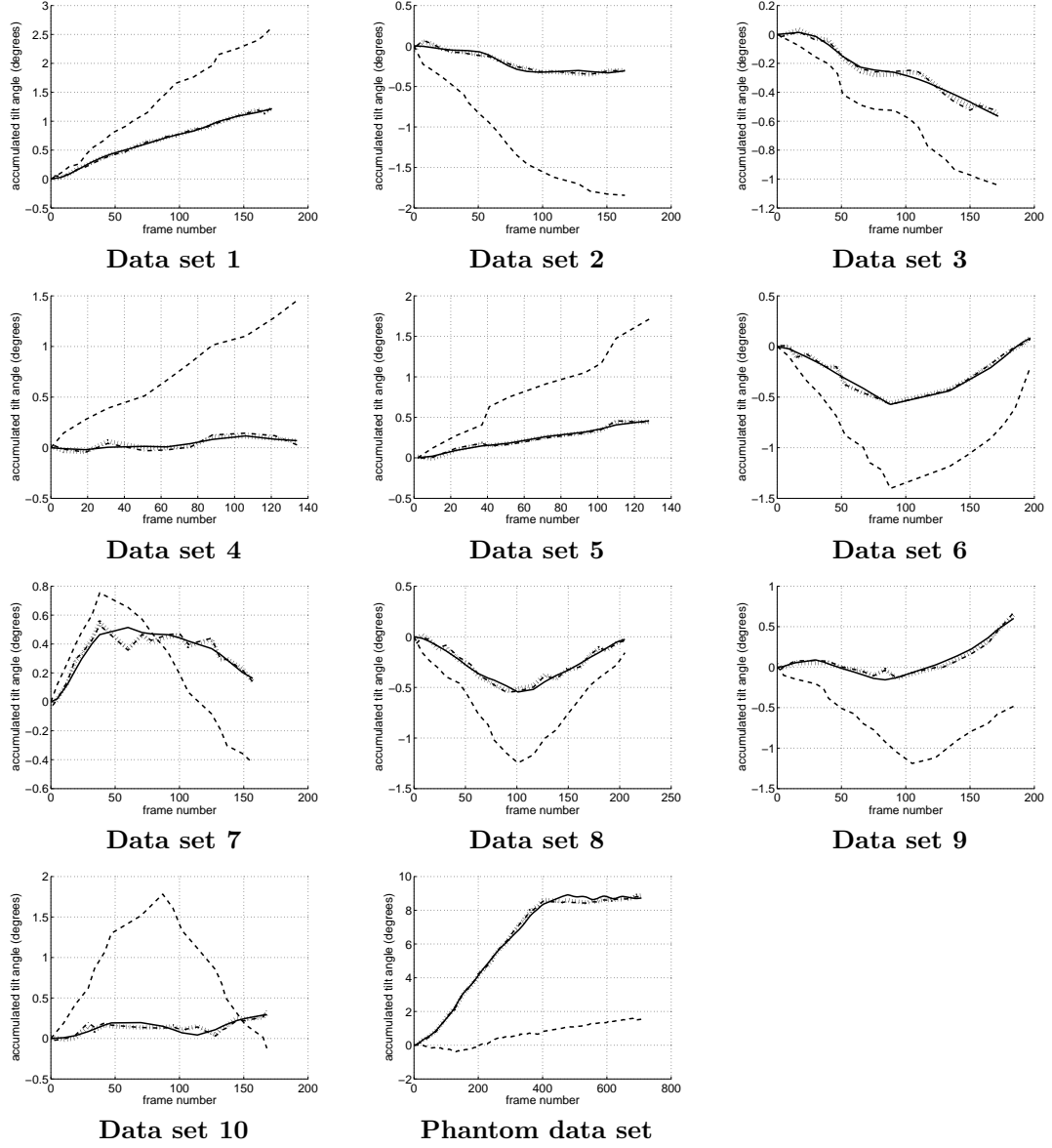
Figure 6: **Drift correction of tilt**. The graphs show the tilt angle accumulated through the data set. The solid lines (—) show smoothed Polaris sensor data. The dashed lines (- -) show the sensorless reconstruction. The six DOF drift-corrected sensorless reconstruction (-·-) and orientation drift correction (|||) are also shown.
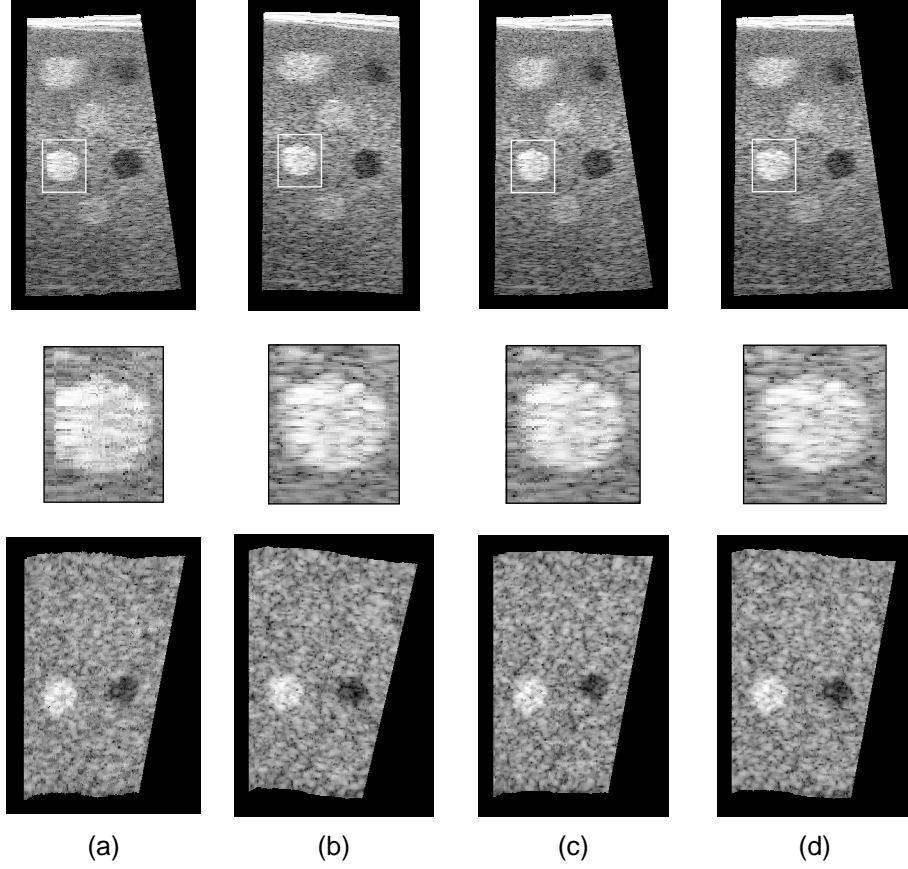
Figure 7: **Reslice images**. The images show reslices through the phantom data set. The top row of reslices are along the length of the data, showing the length and tilt. The central row shows enlarged versions of the outlined region, highlighting the small scale accuracy. The lower reslices are parallel to the scanning surface and therefore show length and yaw. All the reslices use frames from the forward sweep only. (a) Position sensor only reconstruction. (b) Sensorless only. (c) Six DOF correction. (d) Orientation only correction.

of the frame, so fixing the orientation does not affect the centre offset. This is consistent with our earlier findings [9] that the bias variation is centred at the focal depth, which in this case is half way down the image.

It is perhaps not surprising that there is little change in the length, when we have only orientation data to work with. There is, however, a noticeable improvement in the tilt bias in all eleven data sets. In all cases, this is as good as that achieved using the full six DOF sensor.

Figure 7 shows images of reslices through the phantom data set for the various reconstruction and drift correction methods. In this example, we can be more confident that the position sensor has good large scale accuracy. The speckle phantom cannot be distorted by probe pressure in the way that the beef joint could be. The sensorless only reconstruction (Fig. 7(b)) shows a significant tilt error, but no qualitatively obvious length error. The small scale accuracy is also noticeably better for the sensorless reconstruction. Both of the drift correction methods are able to significantly reduce the tilt error, without losing the improved small scale accuracy.
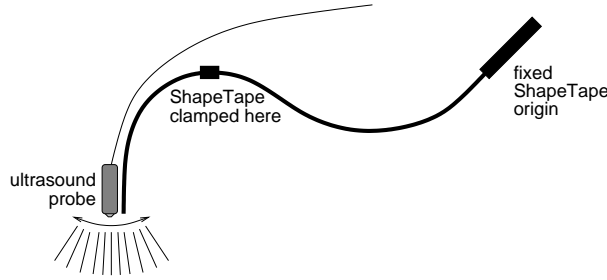
Figure 8: **Experimental setup for the ShapeTape**.

3 Novel position sensors

In the previous section, we described how a position sensor could be used to reduce the drift in an image-based reconstruction and we used a standard optical tracking system to demonstrate its effectiveness. However, we discussed in the introduction how position sensors can be inconvenient to the user. For example, an optical system like the Polaris requires a direct line of sight between the sensor and its origin.

Here we shall describe two alternative sensors that are potentially less obtrusive in the scanning process. The first is a Measurand ShapeTape. It consists of a tape about one metre long with optical fibre sensors along its length. As the tape bends and twists, the optical fibres lose light according to their curvature. The curvature covered by each sensor can be measured and from this the overall shape of the tape can be reconstructed. In our application, the end of the tape is attached to the probe, as shown in Figure 8, so we are only interested in the position and orientation of this point.

There are some potential difficulties with this system. The process of finding the end-point position involves summing up the curvatures from each sensor along the tape and integrating to get the positions. Any small errors in the measured curvatures at each point will be summed up to give a similarly small error in the end-point orientation. However, a small curvature error in a sensor near to the start of the tape will be extrapolated to a large position error at the end, so it is fundamentally difficult for the ShapeTape to accurately measure its end-point position.

Clearly, the accuracy is approximately related to the length of the tape. Our experience with the ShapeTape is that it is best to clamp it close to the end (see Figure 8) so that no more of the length than necessary is free to move. Although there will still be large position errors from sensors near the start of the tape, these will not change during the scan. The relative error in the recorded positions will be limited to that expected from the reduced length.

The ShapeTape is designed to be able to move into any position and orientation and provide a correct measurement. However, we have observed that it is most accurate when the movement is restricted to a single plane, as shown in Figure 8. Where the ShapeTape has been used in other contexts, it has also been found that it is accurate when used in a well constrained manner [15], but has very large errors when allowed to move freely [1]. We therefore restrict our scanning protocol to that shown in the figure. In terms of probe motion, this allows tilted sweeps with mainly elevational motion over several centimetres, but no deliberate in-plane motion. This motion is sufficient for most practical scanning situations, which often use only linear or tilting sweeps of the probe, although it has a very limited operating volume.

The second sensor we consider is an Xsens MT9-B. This device uses rate-of-turn sensors to measure angular motion. This measured data is then integrated to provide orientations. Typically, integrating leads to a gradually increasing error, but the device has addition sensors to correct for this. Magnetometers track the Earth's magnetic field and accelerometers track gravity, so it is able to correct the orientation using this fixed reference.

The most significant limitation we have encountered while using this sensor is the effect of magnetic field distortions. Any ferromagnetic material near the device affects the magnetometer
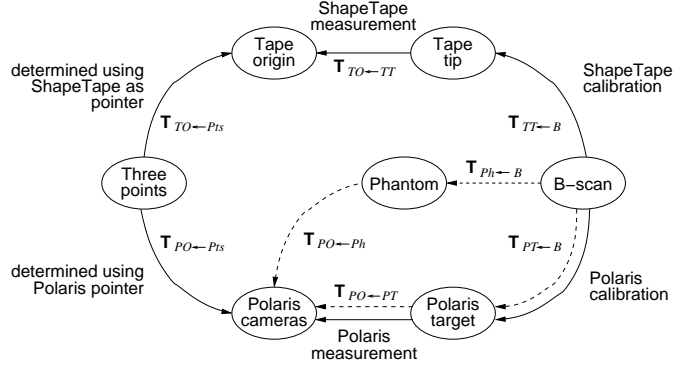
Figure 9: **The ShapeTape calibration process**. The coordinate transformation between the tape tip and the B-scan is found by concatenating five other coordinate transformations. The dashed line shows a standard calibration procedure that could be used for the Polaris system, the solid line shows the ShapeTape calibration procedure.

readings resulting in incorrect orientation measurements. A typical working area is unlikely to be completely free of magnetic field distortions, so care must be taken to avoid them.

3.1 Calibration

Calibration is the process of finding the transformation between the coordinate system of the sensor attached to the probe and the image coordinate system. The sensor measures the location of the sensor coordinate system in world coordinates, but we actually need the image location. We therefore need to know the fixed transformation between the sensor and image so that we can convert between them. Calibration has been studied in detail and a review of current methods is given in [17]. Here, we are dealing with unusual position sensors, where the standard methods are not necessarily suitable.

The ShapeTape is very restricted in its motion: we have discussed how we get the best accuracy by only allowing it to bend in one plane. An approach that attempts to solve for the calibration by moving the ShapeTape through all possible degrees of freedom is therefore not appropriate. We are left with the various techniques that calibrate using just a single position measurement. These techniques involve concatenating the known transformations between several other coordinate systems, where the overall transformation is equivalent to the unknown calibration.

The dashed line in Figure 9 illustrates this for the Polaris system. First, we align the ultrasound image to a known location in a phantom. With knowledge of the phantom geometry, this gives the transformation between the phantom and the image frame, $\mathbf{T}_{Ph \leftarrow B}$. We then measure the location of the phantom relative to the sensor origin, $\mathbf{T}_{PO \leftarrow Ph}$. Finally, the sensor measurement gives the transformation between the origin and the sensor, $\mathbf{T}_{PT \leftarrow PO}$. By concatenating these three known coordinate transformations, we can determine the unknown calibration transformation:

$$\mathbf{T}_{PT \leftarrow B} = \mathbf{T}_{PT \leftarrow PO}\mathbf{T}_{PO \leftarrow Ph}\mathbf{T}_{Ph \leftarrow B}$$

where the various transformations are defined in the figure.

This three transformation approach is not possible with the ShapeTape. The main difficulty is locating the phantom relative to the sensor origin (the start of the ShapeTape). Usually, a phantom is located in sensor origin coordinates by measuring its location with another sensor in the same sensor system, but this is obviously not an option with the ShapeTape. Instead, we concatenate a different set of transformations, shown by the solid line in Figure 9:

$$\mathbf{T}_{TT \leftarrow B} = \mathbf{T}_{TT \leftarrow TO}\mathbf{T}_{TO \leftarrow Pts}\mathbf{T}_{Pts \leftarrow PO}\mathbf{T}_{PO \leftarrow PT}\mathbf{T}_{PT \leftarrow B}$$
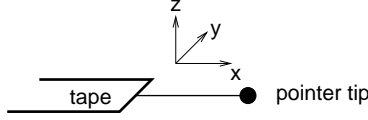
11

Figure 10: **ShapeTape pointer calibration**. The ShapeTape can be used as a pointing device by attaching a narrow rod, with a well defined tip, to the end of the tape. Pointer calibration is usually able to solve for all three components of the translation from the sensor to the pointer tip. Due to the movement restrictions on the ShapeTape, we only solve for the x component and assume the others are 0.

By making use of another sensor system, we bypass the need to find the phantom to tape origin transformation. This alternative method requires that we attach both the ShapeTape and the Polaris targets to the probe at the same time and record a position from each system. This gives us the $\mathbf{T}_{TT \leftarrow TO}$ and $\mathbf{T}_{PO \leftarrow PT}$ transformations. Provided the Polaris system is already calibrated (by any standard method) we also have the $\mathbf{T}_{PT \leftarrow B}$ transformation.

All that remains then is to determine the relative location of the two sensor origins. To do this, we fix three points in space, in the bend plane of the ShapeTape. We can easily locate these points relative to the Polaris origin using a standard pointing device attached to that system. We then use the ShapeTape as a pointer (see Figure 10) to locate the same three points relative to the ShapeTape origin. This gives us the $\mathbf{T}_{TO \leftarrow Pts}$ and $\mathbf{T}_{Pts \leftarrow PO}$ transformations.

One final complication is how to calibrate the pointing devices. The Polaris pointer is calibrated using the standard method [5]. This involves placing the pointer tip at a fixed point and rotating the pointing device around it. This covers all the necessary degrees of freedom to determine the three components of the translation from the sensing device to the pointer tip (see Figure 10). Unfortunately, it is not so simple for the ShapeTape, which cannot be rotated into any position without becoming inaccurate. We therefore have to accept an approximate solution. By rotating it only in the one plane shown in Figure 8, we have enough information to solve for the x component of the translation. The translations in the other two directions, y and z, are assumed to be zero, placing the pointer tip at the centre of the tape's cross section. This is a reasonable assumption given the construction of the ShapeTape.

The process of calibrating the MT9-B is much simpler. Unlike the ShapeTape, there is no movement restriction, so we are able to use our standard plane-based calibration technique [20]. This involves recording several images of a flat plane at the bottom of a water bath. The various images of the plane consist of a straight line through the image, which gives some information about the relative position and orientation of the image frame and calibration plane. Combining this with the recorded positions of the sensor, it is possible to solve for the location of the calibration plane (defined by three parameters) and the transformation between the sensor and the image (six parameters). Solving for these nine parameters requires a fairly complicated scanning protocol to cover the necessary degrees of freedom [21].

For this sensor, we are only looking at orientations, which is a much simpler problem to solve. Rather than using full six DOF coordinate transformations, we only need to consider three DOF coordinate rotations. We can solve for the calibration parameters using a simplified version of the method for a full six DOF calibration. There are now only two parameters defining the orientation of the calibration plane and three for the desired transformation between the image and the sensor. The scanning protocol illustrated in Figure 11 is sufficient to solve for these five parameters.

Rather than using a plane, we use our *Cambridge phantom* [20], which simulates a plane and makes it easier to automatically detect the lines in the images. We also improve the accuracy by taking into account the effect of water temperature on the speed of sound. The speed of sound affects the directions of the lines in the images, so it is important to get this right.

Given only sensor orientation data, there is one ambiguity in the calibration solution. It is possible to rotate the frame 180 degrees in-plane without changing the orientation of the lines in the images. The calibration solution could therefore have a 180 degree in-plane rotation error,
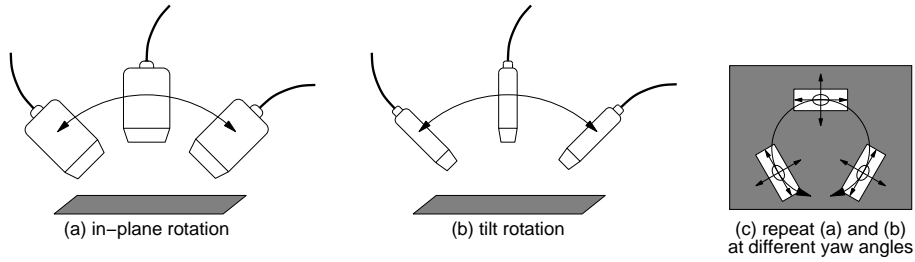
12

Figure 11: **Plane-based calibration protocol for orientation**.

	mean 3D error (mm)	standard deviation (mm)		
		x	y	z
top left	5.17	3.95	2.12	3.45
top right	5.08	3.93	1.86	3.45
centre	5.88	4.80	1.78	3.90
bottom left	6.88	5.68	2.11	4.46
bottom right	6.84	5.66	1.86	4.55

Table 1: **ShapeTape calibration precision**. The statistics shown are the mean 3D error and the standard deviation in each direction, at the frame's centre and corner points. The 3D error for a B-scan point $\mathbf{p}$ is defined as $\|\mathbf{p} - \bar{\mathbf{p}}\|$, where $\bar{\mathbf{p}}$ is the average point location over the 25 calibrations. The standard deviation directions are in the coordinate system of the average frame location.

as this is indistinguishable from the correct solution. Normally, this would be resolved by the additional information in the measured translations, but these are not available here.

We overcome this problem by making use of the known geometry of the experimental setup. Unlike most position sensor systems, the coordinates defined by the sensor origin are not arbitrary. The x direction is aligned with magnetic north and the z direction is upwards, defined by gravity. We also know that the calibration plane is roughly horizontal and viewed from above. On average over all the calibration image frames, we can expect them to be pointing roughly downwards. The average image y direction (defined down the image) should therefore have a negative z component in world coordinates. By checking whether this is the case, we can resolve the ambiguity and correct the calibration if necessary.

3.2   Calibration precision

The precision of the ShapeTape calibration was determined by repeating the calibration process 25 times. Both the ShapeTape and Polaris targets were left attached to the probe between calibrations. The three points used for transforming to the Polaris origin were standardised by fixing them on a mechanical construction. This same set of points was reused for each calibration, but was repositioned approximately in the ShapeTape's bend plane each time. The pointer calibrations for both the ShapeTape and Polaris pointer and the calibration of the Polaris target to the image were done before the calibration experiments and reused each time.

Table 1 shows statistics for the calibration precision. Rather than considering the individual calibration parameters, we have chosen to show the precision in terms of its effect on the location of the image frame. The values show how the location of the frame's corner points and centre point vary.

The precision of the MT9-B calibration was determined in the same way, by repeating the process 25 times. For each calibration, the probe was remounted on the Cambridge phantom and the ultrasound machine gain settings were reset to values within a sensible range for the calibration

	roll	yaw	tilt
standard deviation (degrees)	0.08	1.09	1.68

Table 2: **MT9-B calibration precision**. The statistics shown are the standard deviation, in degrees, of the in-plane roll, and out-of-plane yaw and tilt angles of the 25 calibrations relative to the average calibration.

process. Table 2 shows precision statistics for in-plane roll, and out-of-plane yaw and tilt angles, again in terms of the image frame coordinates.

The ShapeTape calibration imprecision can be almost entirely attributed to the sensor itself, as there was no image or phantom used in the calibration. The only point where a phantom was imaged was to calibrate the Polaris system and this was done using our most accurate calibration system [8], which is considerably more precise than the results shown here. The results show worse precision towards the bottom of the image frame. This is because the tape tip is attached close to the top centre of the frame, so the sensor orientation errors have more effect at the bottom of the image, a larger distance from the sensor.

The precision of the MT9-B can be attributed to the accuracy of detecting the lines in the calibration images. In particular, the tilt is almost entirely dependent on accurately detecting the angles of the image lines during motion (a) in Figure 11. Unfortunately, the line orientation is difficult to measure in this position. The plane is imaged at a glancing angle and so does not produce such a clear image. Also, the temperature correction affects the measured angle of the line. It is therefore not surprising that the worst precision is in the tilt angle.

3.3 Reconstruction accuracy

The following experiments are designed to show the accuracy of the sensors in terms of their application to drift correction. For the ShapeTape, we are interested in how accurately it can reconstruct the length and tilt of an object. For the MT9-B, we need to know how accurately it can measure the relative angles of frames in a sequence. In these experiments, the position of the scanning subject in world coordinates is not known. The accuracy results presented below do not show the accuracy of locating points relative to a fixed reference frame, which is what would be required for an image guided surgery application. They only show the accuracy of locating points relative to each other. This is relevant for applications where measurements are taken from the data, e.g. distances or volumes, without needing to register to an external system.

ShapeTape reconstruction accuracy was determined by recording 25 sweeps of a speckle phantom containing an array of accurately positioned spheres. The spheres were located in the data by identifying the point where each image intersects a sphere. As each sphere was visible in several images, we have several points defining the location of the sphere. We apply a calibration to convert these points to world coordinates (defined by the ShapeTape origin) and then average them to give a final estimate of the sphere position.

The reconstructed grid of spheres was then compared to the known dimensions of the array of spheres in the phantom. The reconstructed grid was registered to the correct grid with a six DOF transformation, i.e. it was translated and rotated, but not scaled. The required transformation was determined by a non-linear optimisation, minimising the sum of squared errors in each of the x, y and z directions. Figure 12 shows two examples of this alignment.

Table 3 shows the position accuracy of individual points. Only the accuracy in the length direction (x in the coordinates of the sphere array) is shown, because the ShapeTape can only be used to correct drift in this one direction. The results are for each of the 25 calibrations, applied to all 25 data sets, resulting in 625 separate reconstructions of the array.

Table 4 shows the accuracy of measuring the length between points and the angle between columns of points in the grid. Length accuracy was determined by considering all possible pairs of points where both points are in the same row of the grid. Tilt accuracy considers all possible pairs of columns of points in the grid, where a column is defined by any two points at the same x
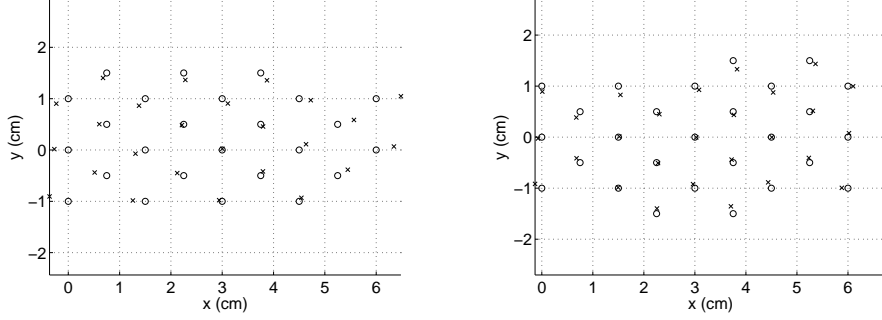
Figure 12: **ShapeTape reconstructions**. The circles show the correct grid and the crosses show the reconstructed grid. The left example is one of the worst results. The right is one of the best.

mean error (mm)	0
mean absolute error (mm)	1.06
standard deviation (mm)	1.41

Table 3: **ShapeTape point location accuracy**. The errors in locating individual points in the direction along the length of the grid (x direction). The mean error is 0 as a result of the least squares alignment of the two grids. For comparison, this statistic is referred to as point location accuracy in [21] and 3D registration accuracy in [16].

position.

The MT9-B reconstruction accuracy was determined by recording 32 sweeps of a Z-wire calibration phantom. This phantom consists of several wires in Z shapes, as shown in Figure 13. The recorded images of the phantom consist of points where the image frames intersect the wires. We use an automatic point detection algorithm to locate these points in the images. The distribution of each set of three points allows us to locate point p_1 in both the B-scan and the phantom. Given at least three of these Z fiducial points, we can locate the image frame in phantom coordinates. The phantom can be imaged from a range of positions, allowing us to record and reconstruct a sweep of frames.

The orientations measured by the sensor were compared to the reconstruction given by the phantom, in terms of roll, yaw and tilt between frames. The two sets of orientations were aligned to minimise the least squares error in each of the roll, yaw and tilt reconstructions. Figure 14 shows examples of the aligned reconstructions.

A complication in this experiment is the precision of the automatic point detection. Ideally, each wire passing through the image frame would result in a circular or eliptical feature in the image and the centre point would be obvious. Unfortunately, due to imperfect focusing of the ultrasound beam and noise in the images, the centre of each wire cannot be located perfectly. An error of a few pixels can be significant, so even with careful automatic point detection, there is noticeable imprecision in the frame location. This causes the jitter that can clearly be seen in Figure 14. In order to overcome this, we apply a smoothing filter to the orientation data from the phantom. All calculations are based on this smoothed data, rather than the original.

As the phantom itself could be a significant source of error, it is worth considering its accuracy for this experiment. We recorded a further data set of the phantom, this time using the Polaris

length error (mm)	-0.29 ± 2.02
tilt error (degrees)	-0.37 ± 2.99

Table 4: **ShapeTape reconstruction accuracy**. The mean and standard deviation of the error in measuring length and tilt between points in the reconstructed sequence.
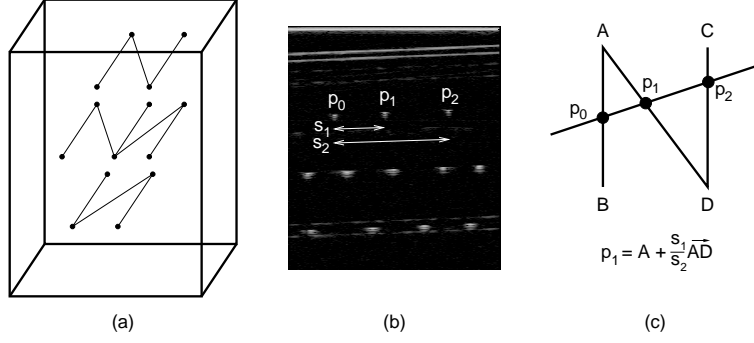
Figure 13: **Z-wire calibration phantom**. (a) The phantom consists of several wires in Z shapes. (b) An image of the phantom consists of points where the image frame intersects the wires. (c) The position of point p_1 can be found in phantom coordinates from measurements s_1 and s_2 taken in the image.

	roll	yaw	tilt
mean error (degrees)	0	0	0
mean absolute error (degrees)	0.22	1.04	0.81
standard deviation (degrees)	0.28	1.32	1.04

Table 5: **MT9-B orientation accuracy**. The error in the orientation of individual frames. The mean error is 0 as a result of the least squares alignment.

sensor. This reconstruction is shown in the bottom row of Figure 14. Looking at these graphs, it is clear that the phantom does not have a large scale drift error, which suggests that the drift in some of the MT9-B results is due to the sensor. The small scale jitter due to imprecise point detection is not a problem, because it is removed by the smoothing filter. However, there are also medium scale errors, for example around frame 35 on the yaw graph and at the start and end of the tilt graph. These are caused by local variations in the appearance of the points in the recorded images. We have been careful to avoid the most obvious of these errors in the MT9-B experiments, by checking the quality of the images.

Table 5 shows the orientation accuracy of individual frames using the MT9-B. The results are derived from the frame orientation errors in all 32 of the data sets, using each of the 25 calibrations, giving 800 separate reconstructions.

For our drift correction application, the most relevant ShapeTape result is the length and tilt accuracy (Table 4). The results show that the ShapeTape will measure lengths and tilts with relatively little bias (the mean value) but with some uncertainty (the standard deviation). In our previous work on sensorless reconstruction [12], we have observed drift errors of around 10% of the length of the sequence. This means that the ShapeTape is likely to improve on the sensorless length drift for sequences more than about 2 cm long. In a typical scan, the recorded sequences will be longer than this, so the ShapeTape will be able to improve the reconstruction in most practical situations.

For the MT9-B experiments, we were careful to avoid the most obvious errors associated with the Z-wire phantom. However, it would be reasonable to attribute some of the observed reconstruction error to uncertainty in the point detection. Given the geometry of the phantom, we would expect point detection errors to have a larger effect on yaw and tilt than on roll and this appears to be the case in Table 5. These results should be taken as an upper bound on the values and we can expect the MT9-B to be more accurate than these values suggest, at least in the yaw and tilt.

The graphs in Figure 6 suggest that there can be large drift errors in the sensorless tilt. For any usefully long sequence, these will give errors of several degrees. Even with the worst case
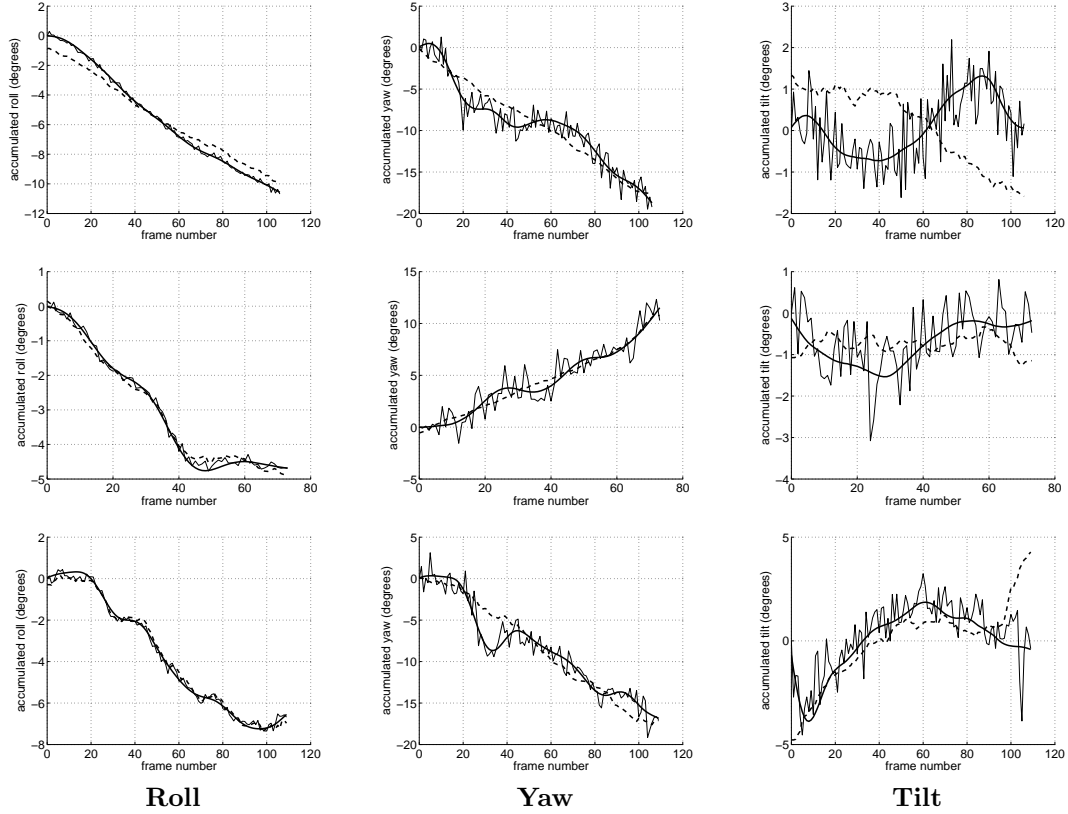
Roll   **Yaw**   **Tilt**

Figure 14: **MT9-B reconstructions**. The graphs show the roll, yaw and tilt measured by the sensor, compared to that given by the Z-wire phantom. The solid lines (—) show the original and smoothed versions of the phantom reconstruction. The dashed line (- -) shows the rotation according to the sensor. The upper row shows one of the worst examples using the MT9-B and the middle row shows one of the best. For comparison, the lower row shows a reconstruction using the Polaris sensor.

accuracy of about 1 degree, the MT9-B is sufficiently accurate to improve on these drift errors. Also, the graphs in Figure 14 suggest that the angle errors vary slowly over the data set rather than as independent random errors. The RANSAC filter described in Section 2 makes use of angles between nearby frame pairs, where the error will be smaller than the overall accuracy in Table 5 suggests.

We have shown in Section 2 that a sensorless reconstruction can be adjusted to match the sensor measurements very closely. This is clearly advantageous when using the Polaris system, since we know that the sensor has good large scale accuracy. For the alternative sensors considered here, we can say that the large scale accuracy of any hybrid system will depend mainly on the sensor accuracy. Provided there is no large scale motion of the scanning subject relative to the sensor origin, the ShapeTape reconstruction accuracy in Table 4 would be a good indication of a hybrid system's performance. We have shown that this is sufficient to improve the accuracy of measurements taken from the data compared to those from a typical sensorless reconstruction. With the MT9-B, it is possible to reduce the orientation drift in a sensorless reconstruction, which will give a partial improvement in the reconstruction accuracy. However, it is not possible to entirely remove the measurement bias using this sensor.

Finally, the purpose of introducing these sensors was as an alternative to the accurate but inconvenient sensors already available, but we have not yet said anything about their convenience. We have shown that the ShapeTape can be accurate enough for our application, but only in a

very restrictive setup. We originally hoped that the ShapeTape could be built into a probe cable so that the user would be unaware of it, but it appears that this is not possible at the accuracy that we require. It would, of course, be possible to use the ShapeTape more flexibly, but at the expense of accuracy. A further limitation is the complexity of the calibration process: it requires the use of another sensor system, which also needs to be calibrated. We have done the calibration in this way because it conveniently makes use of existing calibration equipment. The requirement for a second sensor system could be avoided by building a phantom specifically designed for the ShapeTape. A Z-wire phantom with three locating points within the ShapeTape's bend plane would be suitable.

The MT9-B is a much less obtrusive system. It has no line of sight requirement and no limited operating volume, and it can be attached to the ultrasound probe without restricting the probe motion. The only restriction we have encountered is the effect of magnetic field distortions, but these can be avoided without much difficulty.

4   Conclusions

We have demonstrated that a hybrid system, using both sensorless techniques and limited information from a position sensor, is able to significantly reduce the drift found in purely sensorless reconstructions. At the same time, the improved small scale accuracy in the sensorless reconstructions is maintained. A full six DOF sensor is able to almost completely remove the drift, using a general warping of the reconstruction. When only the orientation data is available, we are still able to achieve a significant improvement in the tilt drift, but there is little improvement in the length.

We have also looked at two alternative position sensors that are potentially less obtrusive than the established sensor systems. We have found that both systems are sufficiently accurate to improve on the sensorless drift. However, since the ShapeTape is both less accurate and more restrictive than the Polaris sensor system, it is difficult to envisage applications where it might serve as a viable alternative. The MT9-B shows more promise in terms of ease of use, but will be less effective for drift correction, given the limited information provided by the sensor.

Acknowledgements

References

[1] Y. Baillot, J. J. Eliason, G. S. Schmidt, J. E. Swan II, D. Brown, S. Julier, M. A. Livingston, and L. Rosenblum. Evaluation of the ShapeTape tracker for wearable, mobile interaction. In *VR '03: Proceedings of the IEEE Virtual Reality 2003*, pages 285–286. IEEE Computer Society, 2003.

[2] L. N. Bohs, B. J. Geiman, M. E. Anderson, S. C. Gebhart, and G. E. Trahey. Speckle tracking for multi-dimensional flow estimation. *Ultrasonics*, 38:369–375, March 2000.

[3] J-F Chen, J. B. Fowlkes, P. L. Carson, and J. M. Rubin. Determination of scan plane motion using speckle decorrelation: theoretical considerations and initial test. *International Journal of Imaging Systems Technology*, 8:38–44, 1997.

[4] L. Danisch, K. Englehart, and A. Trivett. Spatially continuous six degree of freedom position and orientation sensor. *Sensor Review*, 19(2):106–112, 1999.

[5] P. R. Detmer, G. Bashein, T. Hodges, K. W. Beach, E. P. Filer, D. H. Burns, and D. E. Strandness Jr. 3D ultrasonic image feature localization based on magnetic scanhead tracking: *in vitro* calibration and validation. *Ultrasound in Medicine and Biology*, 20(9):923–936, 1994.

[6] A. Fenster, D. B. Downey, and H. N. Cardinal. Three-dimensional ultrasound imaging. *Physics in Medicine and Biology*, 46:R67–R99, 2001.

[7] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

[8] A. H. Gee, N. E. Houghton, G. M. Treece, and R. W. Prager. 3D ultrasound probe calibration without a position sensor. Technical Report CUED/F-INFENG/TR 488, Cambridge University Department of Engineering, September 2004.

[9] A. H. Gee, R. J. Housden, P. Hassenpflug, G. M. Treece, and R. W. Prager. Sensorless freehand 3D ultrasound in real tissue: Speckle decorrelation without fully developed speckle. *Medical Image Analysis*, 10(2):137–149, 2006.

[10] A. H. Gee, R. W. Prager, G. M. Treece, and L. H. Berman. Engineering a freehand 3D ultrasound system. *Pattern Recognition Letters*, 24:757–777, 2003.

[11] P. Hassenpflug, R. W. Prager, G. M. Treece, and A. H. Gee. Distance measurement for sensorless 3D US. In *Proceedings of the 7th International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 1087–1088. LNCS, 2004.

[12] R. J. Housden, A. H. Gee, G. M. Treece, and R. W. Prager. Sensorless reconstruction of unconstrained freehand 3D ultrasound data. *Ultrasound in Medicine and Biology*, in press.

[13] R. J. Housden, A. H. Gee, G. M. Treece, and R. W. Prager. Subsample interpolation strategies for sensorless freehand 3D ultrasound. *Ultrasound in Medicine and Biology*, 32(12):1897–1904, 2006.

[14] M. Li. System and method for 3-D medical imaging using 2-D scan data. United States patent 5,582,173, application number 529778, September 1995.

[15] Y. Li, R. Aissaoui, M. Lacoste, and J. Dansereau. Development and evaluation of a new body-seat interface shape measurement system. *IEEE Transactions on Biomedical Engineering*, 51(11):2040–2050, November 2004.

[16] F. Lindseth, G. A. Tangen, T. Langø, and J. Bang. Probe calibration for freehand 3-D ultrasound. *Ultrasound in Medicine and Biology*, 29(11):1607–1623, November 2003.

[17] L. Mercier, T. Langø, F. Lindseth, and L. D. Collins. A review of calibration techniques for freehand 3-D ultrasound systems. *Ultrasound in Medicine and Biology*, 31(2):143–165, February 2005.

[18] J. J. More. The Levenberg-Marquardt algorithm: implementation and theory. In A. Watson, editor, *Numerical Analysis*, pages 105–116. Lecture Notes in Mathematics 630, Springer-Verlag, 1977.

[19] T. R. Nelson and D. H. Pretorius. Three-dimensional ultrasound imaging. *Ultrasound in Medicine and Biology*, 24(9):1243–1270, 1998.

[20] R. W. Prager, R. N. Rohling, A. H. Gee, and L. Berman. Rapid calibration for 3-D freehand ultrasound. *Ultrasound in Medicine and Biology*, 24(6):855–869, July 1998.

[21] G. M. Treece, A. H. Gee, R. W. Prager, C. J. C. Cash, and L. Berman. High definition freehand 3D ultrasound. *Ultrasound in Medicine and Biology*, 29(4):529–546, April 2003.

[22] G. M. Treece, R. W. Prager, A. H. Gee, and L. Berman. Correction of probe pressure artifacts in freehand 3D ultrasound. *Medical Image Analysis*, 6(3):199–215, 2002.

[23] T. A. Tuthill, J. F. Krücker, J. B. Fowlkes, and P. L. Carson. Automated three-dimensional US frame positioning computed from elevational speckle decorrelation. *Radiology*, 209:575–582, 1998.

[24] L. Weng, A. P. Tirumalai, C. M. Lowery, L. F. Nock, D. E. Gustafson, P. L. V. Behren, and J. H. Kim. US extended-field-of-view imaging technology. *Radiology*, 203(3):427–441, 1997.