

Sensorless Reconstruction of Freehand 3D Ultrasound Data

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

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

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. Previous work has not considered how to reconstruct entirely freehand data, which can exhibit irregularly spaced frames, non-monotonic out-of-plane probe motion and significant in-plane motion. This paper presents reconstruction methods that overcome these limitations and are able to robustly reconstruct freehand data. The methods are assessed on freehand data sets and compared to reconstructions obtained using a position sensor.

1 Introduction

Freehand 3D ultrasound is a technique that produces 3D data from a series of 2D ultrasound images [1]. In order to do this, the position of each individual scan must be known relative to the others. This is normally achieved using a position sensor, but the extra sensor is an inconvenience. A more serious limitation is the small scale accuracy of the sensor — over a short range, the positioning errors can be significant. Also, a position sensor gives positions relative to a fixed reference, which relies on the scanning subject remaining stationary during the scan.

An alternative is image-based positioning, which uses information in the images themselves. The offset in the plane of the images can be found using standard image registration and motion tracking techniques [2,3]. Out of the plane of the image, speckle decorrelation [4,5,6] can be used to estimate the absolute separation between two patches of data. The idea is that the correlation between the patches varies with their separation. By calculating the correlation, the value can be converted to a distance using a precalibrated decorrelation curve.

There are several complications that must be overcome in order to apply this basic theory to a typical freehand data set. First, the calibrated decorrelation curve is only valid for data consisting of fully developed speckle, but there is very little of this in scans of real tissue. A solution to this problem is to use the adaptive speckle decorrelation method presented in [7]. Secondly, there are other causes of speckle decorrelation that might be confused for elevational motion. While most are a matter for future research, the effects of in-plane motion have been studied in detail. In [8], it is shown how to interpolate the in-plane motion estimates to sub-sample precision and then correct the elevational correlation values by allowing for the sub-sample in-plane motion.

In addition, the speckle decorrelation method only gives the absolute value of the separation between two patches. Previous work has ignored this problem and instead constrained the motion to be monotonic with no intersecting frames, in which case all the offset estimates are in the same direction. In this paper, we present for the first time sensorless reconstructions of truly unconstrained freehand data. Building on preliminary work [9] with regularly spaced frames, we present algorithms to allow for arbitrary hand-held probe motion, including cases where the probe reverses direction part-way through the scan.

The process of reconstructing a freehand 3D data set is as follows:

1. Each image frame is divided into a grid of 12×8 patches.
2. For each frame A , we determine the set of frames B_i within range of A . By “within range”, we mean that the majority of the patches on B_i are offset by an amount within the search range of the in-plane alignment and are close enough elevationally to correlate to some extent with a corresponding patch on A . Frames that are elevationally too distant from A will be totally decorrelated and not allow any meaningful estimation of elevational or in-plane separation.
3. For all frame pairs within range of each other, the offset between corresponding patches is estimated as three translational degrees of freedom (two in-plane and one out-of-plane). This is done as accurately as possible, making use of the methods in [7,8].
4. Using this elevational separation information, several interleaved sets of coarsely spaced frames are selected from the irregularly spaced sequence of frames, using the method described in Section 2.
5. The coarse sets of frames are reconstructed independently and then combined into a final reconstruction, as described in Section 3.

It is assumed that most of the frame pairs will not intersect at the coarse scale, so there is no need to correct for intersections. This is a reasonable assumption for linear scanning protocols. Other scanning protocols are also possible, for example axial rotation, where there will certainly be intersections. Reconstructing these data sets is the subject of related work [10].

2 Frame Selection

A well known problem in sensorless reconstruction is drift errors from accumulating offsets through a sequence of frames. They are caused by biases in the elevational separation estimates which are most significant when the frames are closely spaced. Our frame selection algorithm therefore has two objectives. First, a coarse set of frames must be widely spaced to minimise drift errors. Also, the multiple coarse sets must use different frames, so that they give independent reconstructions. It is not always possible to achieve both of these objectives, in which case the independence objective is considered to be more important.

Consider initially the problem of choosing only one set of frames from the sequence, so that independence is not an issue. The algorithm builds up a coarse

set of frames one frame at a time, using range information from the two previous frames in the coarse sequence. Consider Figure 1, where frames A and B are already chosen and the next frame C is to be found. Clearly, C must lie within range of B , giving the $r_{\max}(B)$ limit. However, in order to detect non-monotonic motion, we need to ensure that each frame is within range of not one but two frames further along the sequence (the reason for this will become clear in Section 3). Ideally, one of these frames would be at the maximum range and the other half way between. C must therefore be positioned within the $r_{\text{half}}(B)$ range and must also be within range of A . The ideal choice for C is therefore the smaller of $A + r_{\max}(A)$ and $B + r_{\text{half}}(B)$.

This method would generate a single suitably spaced sequence. However, with multiple coarse sequences, it is possible that a frame would be the ideal next frame for more than one sequence. It is then necessary to adjust the frames used by each set to attempt to avoid repetitions. Figure 2 shows an example of this situation. Three sets of frames are being constructed together. We have reached a point where two coarse sets (1 and 3) both have the same ideal next frame,

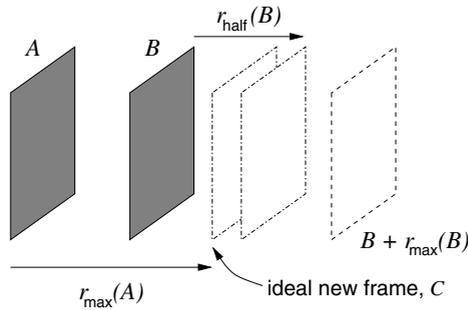


Fig. 1. Choosing the next frame in the sequence. The ideal next frame depends on two maximum limits: it must be within range of both the preceding frames, A and B , and leave room for another frame within range of B .

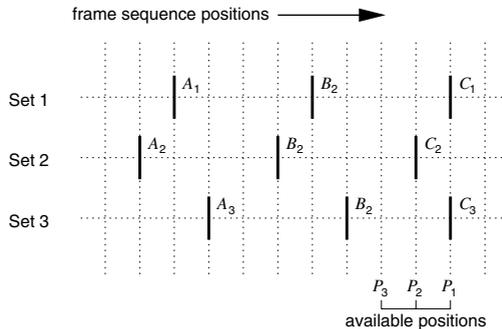


Fig. 2. Repeated frame in coarse sequences. Sets 1 and 3 both expect the same frame as their next frame. The algorithm adjusts them if there is room to do so.

so we need to try to adjust the selections. Since we wish to change no more than the last frame in each set, the only slots available are those after frame B_3 , labelled P_1 to P_3 in the figure. There are two frames to adjust, so we initially consider just P_1 and P_2 . However, as frame C_2 is already in position P_2 , we find that we actually need to adjust all three frames, using all three available slots. In more extreme cases, there might not be enough slots to play with. We must then accept a repeated frame as unavoidable, but continue to adjust any further repetitions. Having established how many frames to adjust, we distribute the frames evenly into the available slots, minimising the number of times each position is used. We order the frames according to the order of the preceding frames in the sets. In our example, this means that both frames C_2 and C_1 are moved back one position, while C_3 is unchanged.

3 Robust Reconstruction

Consider reconstructing a sequence of frames. The out-of-plane offset between two frames can be found by solving a simple least squares problem $\mathbf{d} = \mathbf{X}\mathbf{w}$, where \mathbf{d} is a vector of patch-wise distance estimates and \mathbf{w} is a vector of three out-of-plane offset parameters (centre elevational offset and gradients across and down the frame). \mathbf{X} is a coefficient matrix.

The direction ambiguity in the elevational distance estimates means that it is not immediately possible to see when the probe reverses direction. Given three frames A , B and C , the AB and BC offsets can be found independently, but their relative directions are not known. By making use of the distances between A and C , an expanded least squares problem can be solved for the three frames together, disambiguating the two possible directions of BC relative to AB :

$$\begin{pmatrix} \mathbf{d}_{AB} \\ \mathbf{d}_{AC} \\ \mathbf{d}_{BC} \end{pmatrix} = \begin{pmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{X} & \mathbf{X} \\ \mathbf{0} & \mathbf{X} \end{pmatrix} \begin{pmatrix} \mathbf{w}_{AB} \\ \mathbf{w}_{BC} \end{pmatrix} \quad (1)$$

Figure 3(a) shows how we use this multiframe approach to reconstruct the coarsely spaced sequences of frames. We consider three consecutive frames in a sequence as an independent block and position them using equation (1). We then consider another block of three frames that overlaps the first, as shown in the figure. These independently positioned blocks are joined together by averaging the offset in the overlapping range. By repeating this process, we can build up an entire sequence of frames.

This can easily give an incorrect reconstruction, as it only needs one spurious change of direction to make the entire sequence useless. The important next step is therefore to exploit the redundancy afforded by the multiple coarse sets (Figure 3(b)). We look for consistency between the sets by finding the maximum difference in their accumulated offsets. Reconstructions that disagree on changes of direction will have a much larger maximum difference than similar reconstructions. Sequences with a difference of less than 0.5 mm are taken as consistent.

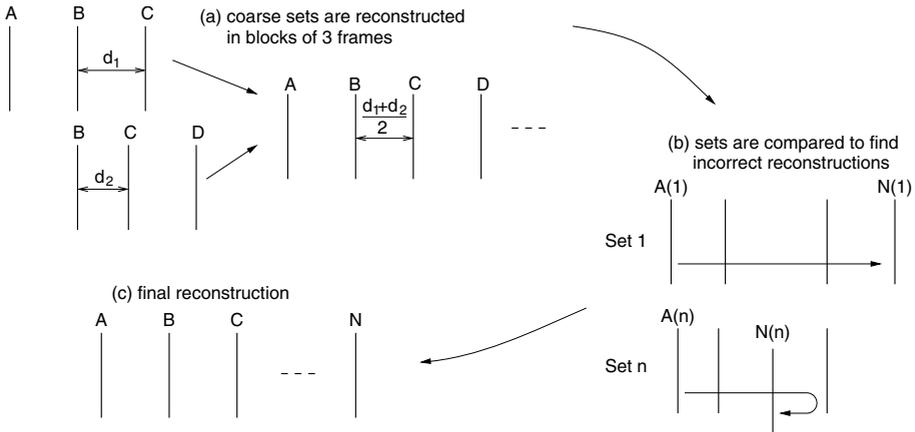


Fig. 3. The reconstruction algorithm

Having rejected the incorrect reconstructions, a final scale is determined by averaging the maximum lengths of the remaining sets. This averaging is intended to reduce any non-systematic drift error. The final reconstruction consists of the frames of one of the sets, scaled to match the average length (Figure 3(c)). Note that there will be one direction ambiguity remaining in the final reconstruction, so it could be the mirror image of the correct reconstruction. There is currently no solution to this problem other than to manually determine which is correct either from the scanning protocol or from features in the images.

4 Experiments and Results

Ultrasound data was recorded with a 5–10 MHz linear array probe connected to a Dynamic Imaging Diasus ultrasound machine (<http://www.dynamicimaging.co.uk>). The depth setting was 4 cm with a single focus at 2 cm. Analogue radio frequency (RF) ultrasound signals were digitised after receive focusing and time-delay compensation, but before log-compression and envelope detection, using a Gage Compuscope CS14200 14-bit digitiser (<http://www.gage-applied.com>). The system works in real time, with acquisition rates of about 30 frames per second. The signal was sampled at 66.67 MHz, synchronous with the ultrasound machine's internal clock. The RF vectors were filtered with a 5–10 MHz filter, then envelope detected using the Hilbert transform. The resulting 127×3818 frames of backscatter amplitude data formed the basis of all further computation: speckle decorrelation and lateral interpolation [8] require RF amplitude data rather than rasterised B-scans. The frame positions were recorded using a Polaris optical tracking system (<http://www.ndigital.com>).

Several 3D data sets were recorded by scanning beef joints by hand in both monotonic and non-monotonic sweeps. The only deliberate motion was elevational, but being freehand, there were also small in-plane and rotational motions between each frame pair. The data had very dense frame spacing, so that the

frame selection algorithm was able to select five independent sequences from each. Ten data sets were considered, five with no deliberate change of direction (although there could be some jitter in the probe motion) and five with the probe deliberately reversed once. Figure 4 shows the reconstruction in terms of the offset at the centre of the frames, accumulated through the data set, for each set of data considered. All five of the coarse sets are shown separately, along with the equivalent values from the position sensor.

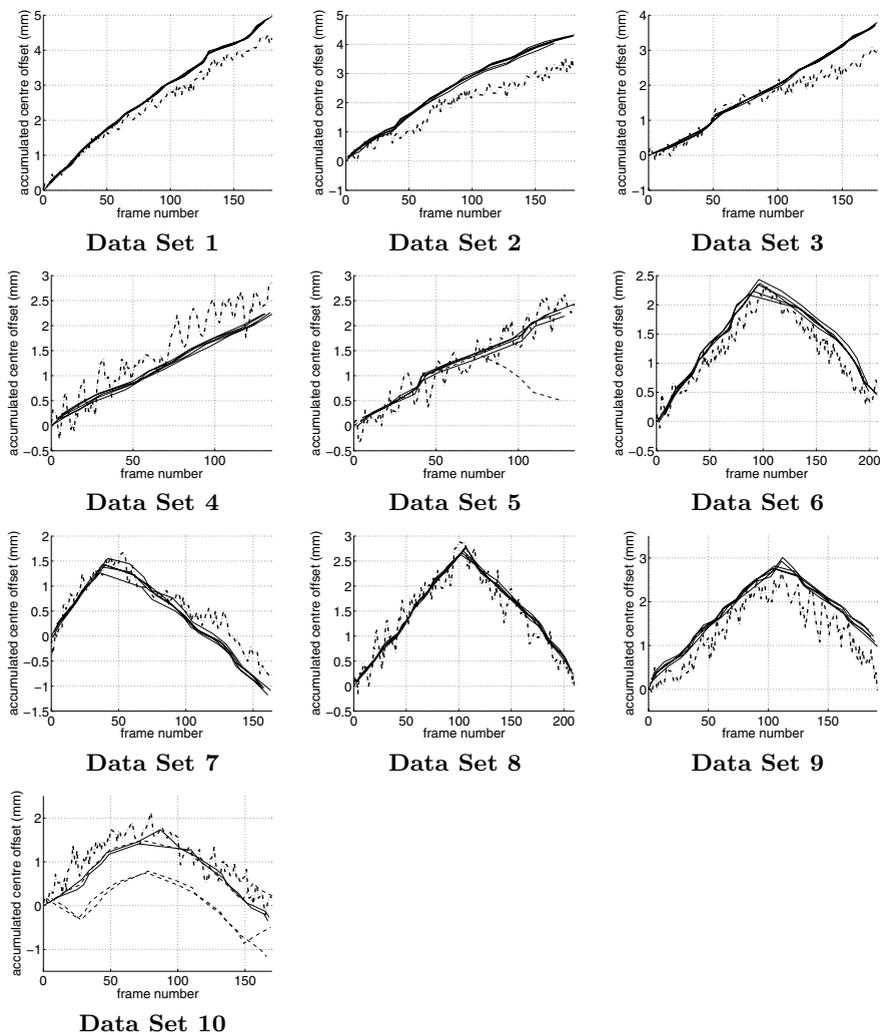


Fig. 4. Reconstruction of freehand data sets. Each figure shows the coarse scale accumulated centre offsets for the five independent sets of frames. The solid lines (—) are the sets that contribute to the consensus. The dashed lines (- -) are those that are rejected. The centre offsets as given by the position sensor are also shown (-.-). The first five data sets are monotonic and the rest have one deliberate turning point.

A meaningful comparison between the two reconstructions would require the beef to remain stationary with respect to the Polaris system's reference frame. To this end, the probe was separated from the beef by a thick layer of coupling gel, to minimise the chance of any contact force displacing the beef. Note that this precaution is solely for the benefit of the sensor-based reconstructions. Sensorless reconstructions are not sensitive to rigid movement of the scanning subject.

The results in Figure 4 demonstrate how the algorithm can robustly reconstruct a data set. In all except the last example there is a clear consensus among the coarse reconstructions, with at most one set different. In addition, the consensus sets give very similar reconstructions, despite using independent distance estimates for most, if not all, of their length. This suggests that uncertainty in the distance estimates does not seriously affect the accuracy of the overall reconstruction. In the worst case, data set 10, the final reconstruction relied on a consensus of two from five. Data set 10 featured two points with uncertain monotonicity, whereas most data sets have none.

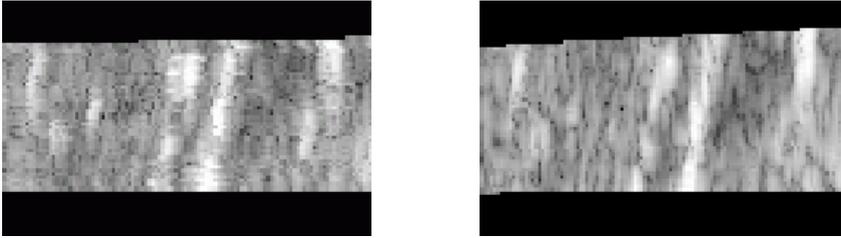


Fig. 5. Reslices. The figure shows a part of a reslice along the length of data set 1. The left reslice is derived from a position sensor reconstruction, whereas the right one is from the image-based reconstruction.

The sensorless reconstructions differ from the position sensor reconstructions in two ways: small scale jitter and large scale drift. The position sensor has significant small scale inaccuracy, which affects the quality of the reslice image shown in Figure 5. In contrast, the sensorless reslice is much smoother.

We have used coarsely spaced frames in order to minimise the large scale drift error, but it is clear from Figure 4 that we have not eliminated it entirely. The lateral interpolation between RF vectors [8] and the adaptive speckle decorrelation [7] both contribute to the bias. A further source of bias is transducer rotation between frames. This would result in systematic overestimation of elevational separation — as is evident in many of the results in Figure 4 — and helps explain the somewhat larger drift for, say, data set 2 compared with previous results obtained with purely translational probe motion [7,8].

5 Conclusions

This paper has presented results showing the viability of sensorless reconstruction of freehand 3D ultrasound data. The method has been shown to be

robust to the complex motions in freehand data, particularly non-monotonic motion. Image-based reconstruction has advantages over the usual position sensor method, particularly for small scale accuracy. Also, it shows movement relative to the scanning subject, which is an advantage if the subject is free to move. However, unavoidable drift limits the accuracy of any large scale measurements based on the sensorless reconstructions.

References

1. Fenster, A., Downey, D.B., Cardinal, H.N.: Three-dimensional ultrasound imaging. *Physics in Medicine and Biology* **46** (2001) R67–R99
2. Weng, L., Tirumalai, A.P., Lowery, C.M., Nock, L.F., Gustafson, D.E., Behren, P.L.V., Kim, J.H.: US extended-field-of-view imaging technology. *Radiology* **203**(3) (1997) 427–441
3. Geiman, B.J., Bohs, L.N., Anderson, M.E., Breit, S.M., Trahey, G.E.: A novel interpolation strategy for estimating subsample speckle motion. *Physics in Medicine and Biology* **45** (2002) 1541–1552
4. Chen, J.F., Fowlkes, J.B., Carson, P.L., Rubin, J.M.: Determination of scan plane motion using speckle decorrelation: theoretical considerations and initial test. *International Journal of Imaging Systems Technology* **8** (1997) 38–44
5. Li, M.: System and method for 3-D medical imaging using 2-D scan data. United States patent 5,582,173, application number 529778 (1995)
6. Tuthill, T.A., Krücker, J.F., Fowlkes, J.B., Carson, P.L.: Automated three-dimensional US frame positioning computed from elevational speckle decorrelation. *Radiology* **209** (1998) 575–582
7. Gee, A.H., Housden, R.J., Hassenpflug, P., Treece, G.M., Prager, R.W.: Sensorless freehand 3D ultrasound in real tissue: Speckle decorrelation without fully developed speckle. *Medical Image Analysis* **10**(2) (2006) 137–149
8. Housden, R.J., Gee, A.H., Treece, G.M., Prager, R.W.: Sub-sample interpolation strategies for sensorless freehand 3D ultrasound. Technical Report CUED/F-INFENG/TR 545, Cambridge University Department of Engineering (2006)
9. Housden, R.J., Gee, A.H., Prager, R.W., Treece, G.M.: Sensorless freehand 3D ultrasound for non-monotonic and intersecting frames. In: *Proceedings of the 9th Annual Conference on Medical Image Understanding and Analysis*, Bristol, UK (2005) 127–130
10. Housden, R.J., Gee, A.H., Treece, G.M., Prager, R.W.: Sensorless reconstruction of unconstrained freehand 3D ultrasound data. Technical Report CUED/F-INFENG/TR 553, Cambridge University Department of Engineering (2006)