A Novel Temporal Calibration Method for 3D Ultrasound

François Rousseau*†‡§,Pierre Hellier*†‡§,Christian Barillot *Senior Member, IEEE*,*†‡§
*CNRS, IRISA, Unit/Project Visages U746, F-35042 Rennes, France
†INRIA, IRISA, Unit/Project Visages U746, F-35042 Rennes, France
‡INSERM, IRISA, Unit/Project Visages U746, F-35042 Rennes, France
§University of Rennes I, IRISA, Unit/Project Visages U746, F-35042 Rennes, France

***Abstract*— This paper examines a novel approach for temporal calibration of a 3D freehand ultrasound (US) system. A localization system fixed on the probe gives the position and orientation of the probe. For quantitative use, calibration is needed to correctly localize a B-scan in 4D (3D+t) space. Temporal latency estimation is defined in a general robust formulation using no specific probe motion constraints. Experiments were performed on synthetic and real data using a 3D freehand ultrasound system. The achieved precision is lower than the image acquisition rate ($40ms$). A validation study using a calibration phantom has been performed to evaluate the influence of incorrect latency estimation on the 3D reconstruction procedure. We showed that for latency estimation errors less than $40ms$, the 3D reconstruction errors are negligible for volume estimation.**

***Index Terms*— 3D Freehand Ultrasound, Temporal Calibration**

I. Introduction

Conventional ultrasound exams are limited by 2D viewing, and follow-up studies are not easily reproducible [1]. 3D ultrasound imaging overcomes these limitations and allows quantitative accurate morphometric studies. One of the most flexible solutions for 3D ultrasound is a 3D freehand system [2]. This solution consists of fixing a localization system on the probe, which continuously gives the position and orientation of the probe [2], [3]. In this paper, we consider thus this type of acquisition system. The localization system can be magnetic, optic, acoustic or mechanical [1]. Knowing positions and orientations of each B-scan, a 3D image can be reconstructed.

Calibration is required to correctly localize a B-scan in 4D (3D+t) space and includes temporal and spatial calibration. The aim of temporal calibration is to match the position information with the echographic frames. Spatial calibration consists of determining the transformation between pixels in the US image and points in the 3D space [3], [4]. Calibration is crucial because it has a significant impact on the quality of the reconstruction. In addition, the calibration procedure has to be as easy as possible to work in a clinical context, i.e., it needs to be automatic, user friendly, easy to operate, robust and fast.

Temporal calibration, which is the focus of this work, consists of estimating the latency between the position sensor timestamps and ultrasound image timestamps. Temporal calibration depends on the experimental setup, including video grabber and PC used, echographic machine and video transfer format, position sensor and transfer rate, image quality of the phantom used for calibration. Therefore, temporal calibration is specific to each experimental setup. Temporal calibration has previously been achieved by estimating the latency using visual criteria for augmented reality [5]. Jacobs et al. [5] have shown that the latency estimation is a crucial problem. Their system is composed of an ultrasound probe held with a mechanical arm and a virtual reality helmet tracked by a magnetic sensor. For images acquired at 10*Hz*, the measured latency was $220ms$.

In freehand 3D ultrasound, the StradX software [3] is considered as one of the state-of-the-art approaches. It encompasses two temporal calibration methods. The principle is to match changes appearing in ultrasound images and changes in position data. The first method is based on break detection while the second method uses parametric regression [6].

Finally, Nakamoto *et al.* [7] proposed a temporal calibration method using a point based phantom. Latency estimation is performed with a least-square minimization of the distance between the phantom point and the estimated position of this point.

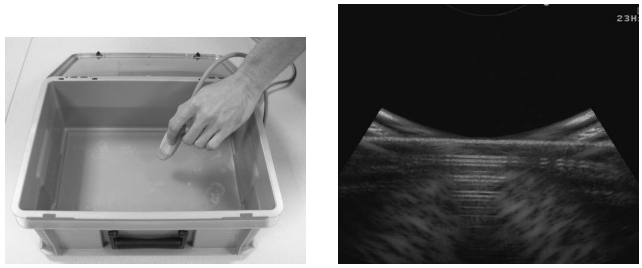
For 3D quantitative studies, temporal calibration is needed if latency is significant. The method suggested by Nakamoto seems to provide a reliable estimate of latency. However, point based phantoms are not easy to use. The plane phantom is attractive because it is very simple to build and easy to use. From our experience, StradX line detection can fail for noisy sequences. The motivation of this work is twofold: firstly, we propose an alternative method to solve the latency estimation problem. Our method uses a robust and automatic line detection algorithm and a general formulation of the latency problem estimation for 3D freehand ultrasound. Secondly, we aim at estimating the impact of temporal miscalibration on 3D quantitative measurements.

II. Method

With the aim of proposing a simple and automatic temporal calibration, we have chosen to use a plane phantom. Figure 1 shows this phantom which consists of a plexiglas plate in a water bath. This type of phantom is easy to build and to use.

François Rousseau is now with Université Louis Pasteur (ULP), 67085 Strasbourg, France, Laboratoire des sciences de l'image, de l'informatique et de la télédétection (LSIIT, UMR CNRS-ULP 7005), 67412 Illkirch, France

Line extraction is performed using an automatic and robust algorithm [4]. A general formulation of latency estimation is proposed in order to handle any image sequences without any probe motion constraints. The basic idea of this work is to fit a parametric model to the position signal and the position of extracted lines. We then estimate a temporal calibration that matches the break detections in these two signals.



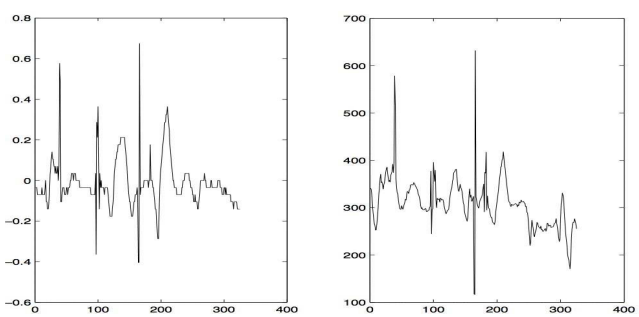
(a) (b)

Fig. 1. In (a), plane phantom used to estimate temporal latency; in (b), typical ultrasound image of the phantom.

A. Line Detection Algorithm

This work is built on our previous work where we presented a reliable line detection method [4]. In each B-scan, we extract points of interest belonging to the line produced by the water bath, using gradient and luminance information. The consistency of the extraction is checked successively using 2D information (Hough transform) and 3D information (spatio-temporal continuity of the point set).

To reject outliers in each image , the Hough transform is used to extract lines in ultrasound images [8]. The estimated line is defined with the following equation: $x\cos\theta + y\sin\theta - \rho = 0$, where ρ is the length of a normal from the origin to this line and θ is the orientation of ρ with respect to the X-axis in the B-scan. A distance criterion between the estimated line and the point set is used to reject possible outliers.



(a) (b)

Fig. 2. Evolution of the line parameters estimated using the Hough transform. (a): $\frac{\cos\theta}{\sin\theta}$ and (b): $\frac{-\rho}{\sin\theta}$.

Continuity of probe motion should lead to a smooth variation of line parameters. Using the assumption of spatio-temporal continuity of the point set, we reject images where the line extraction is not correct. We use a parametrization of the line using a couple of parameters $(\frac{\cos\theta}{\sin\theta}, \frac{-\rho}{\sin\theta})$. Figure 2 plots the couple evolution for a sequence. First, B-spline approximation of the evolution of line parameters is performed. Let $\mathcal{T} = \{\tau_i, i = 1\ldots, n\}$ denote the knot set. A m order B-spline is defined as [9]:

$$B_i(t) = (\tau_{i+m+2} - \tau_i) \sum_{j=0}^{m+2} \left(\frac{(\tau_{i+j} - t)_+^{m+1}}{\prod_{l=0:l\neq j}^{m+2} (\tau_{i+j} - \tau_{i+l})} \right)$$

with

$$(\tau_{i+j} - t)_+^{m+1} = \begin{cases} (\tau_{i+j} - t)^{m+1} & if \quad t \leq \tau \\ 0 & if \quad t > \tau. \end{cases}$$

The approximation model is the following: $Y_i = g(t_i) + W_i,\ with\ g(t_i) = \sum_{j=0}^{q+1} \beta_j B_j(t_i)$ where Y_i is the raw signal, g is the regression spline, W_i are independent random variables and B_j are the B-splines. In our case, we used cubic B-splines ($m = 3$) with equidistant knots. Then, in order to reject line outliers, a statistical test is performed on the differences between these approximations and line parameter evolutions.

This line detection algorithm has been used previously for spatial calibration [4] and has been shown to be reliable and robust.

B. Latency Estimation

The latency is estimated by matching the position sensor signals with the line parameters. Figures 2 and 3 show evolutions of line parameters and positions provided by the localization system throughout an image sequence. Using a plane phantom, some degenerate motion is possible, i.e. motion that will not lead to a variation of the plane seen in the images. For instance, a rotation around the probe depth axe will cause little change in the images. However, this assertion is not symmetric: a significant change of the line position in the image surely corresponds to a variation of the probe motion. We will exploit this asymmetry and match line position ruptures with changes in the signal position.

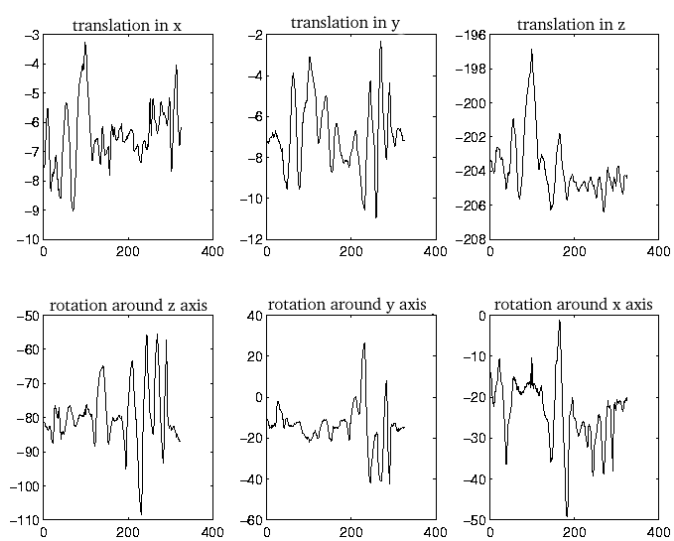


Fig. 3. The position sensor signal (3 translations and 3 rotations).

We detect changes in both signals (image signals and position signals) and match them. For the image signals, two features are available: the slope and the intercept. We experienced that the intercept signal $\frac{-\rho}{\sin\theta}$ is less noisy and variations are larger than for the signal $\frac{\cos\theta}{\sin\theta}$ related to the line slope. Therefore we will retain the intercept and denote $sr = \frac{-\rho}{\sin\theta}$ the "reference signal". The latency estimation is carried out by matching extrema of the position signals with the extrema of the reference signal.

Let sr denote the reference signal and sp_i the position sensor signals, $i \in \{1, \ldots, 6\}$ (3 translations and 3 rotations). Let $\mathcal{G}$ be the point set for which gradient of sr is zero: $\mathcal{G} = \{x, \frac{dsr}{dt}(x) = 0\}$. Latency $\mathcal{L}$ is estimated so as to minimize the sum of gradient for the points in $\mathcal{G}$. We thus minimize:

$$\mathcal{L} = \arg\min_u \sum_{x \in \mathcal{G}} \sum_{i=1}^{6} \left(\frac{dsp_i}{dt}(x+u) \right)^2 \qquad (1)$$

The position sensor signals and reference signal are noisy. A Gaussian filtering is applied to remove high frequency variations (standard deviation $\sigma = 1$). Then each signal is approximated using B-splines. Accurate gradient estimation is performed using the analytical formulation of every signal obtained from B-spline parametrization. Figure 4 illustrates the filtering procedure and the zero gradient point extraction for the reference signal. The continuous parametrization of the signals allows the estimated latency to be lower than the image acquisition rate.

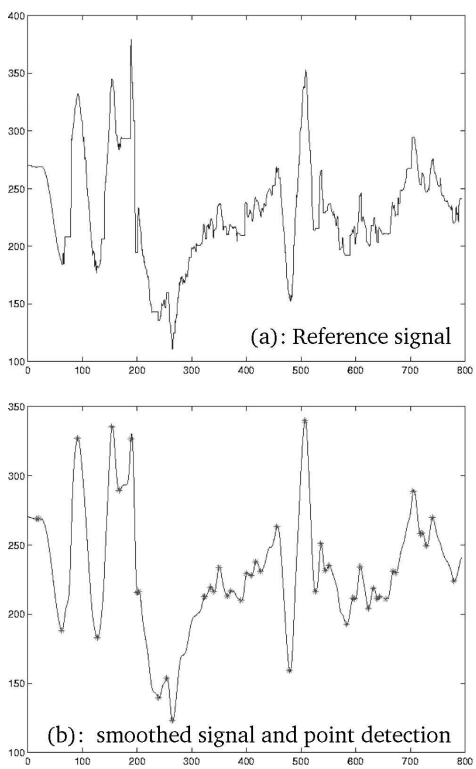


Fig. 4. Reference signal smoothing and zero gradient point extraction. (a): raw reference signal, (b): smoothed reference signal and break detection. We manage to detect accurately the major (or relevant) signal changes.

As outliers may perturb the latency estimation, robust estimators are introduced in order to reject points with high gradient value for all position signals. The robust formulation takes into account only the position signals which take part into the visible changes in the image and thus in the reference signal variations. We chose to use M-estimators for their proven effectiveness. Let $y = \{y_i, 1 \leq i \leq m\}$ be the data set with $y_i = f(\theta, x_i) + r_i$. Parameters θ are estimated taking into account the data x_i and the noise r_i. The principle of these estimators consists of minimizing the sum of the residuals: $\hat{\theta} = \arg\min_\theta E(\theta)$ with $E(\theta) = \sum_i \rho(y_i - f(\theta, x_i))$. The function ρ is called M-estimator. Additional details about properties of M-estimators can be found in [10]. Six weights are therefore attributed to each point (one weight for each position signal). In practice, the parameter of the robust function σ is fixed arbitrarily at 5. We experienced that the method leads to comparable results for a large range of variation of parameter σ.

In addition, we introduce a confidence measure for the points of interest. Basically, we aim at retaining only peaks, i.e. points that are significantly informative. This is motivated by the complexity of the image sequences (noise and low frame rate). A quality measure of the points of interest is proposed to improve the robustness and is defined as:

$$q(x) = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{dsr}{dt}(x) - \frac{dsr}{dt}(x_i) \right)^2$$

where x_i is a neighbor of x, n is the number of neighbors considered (typically, $n = 6$). The larger $q(x)$, the more interesting x. A small $q(x)$ corresponds to a flat area which is not relevant for matching. The zero gradient point selection is carried out by retaining only points for which the q value is large, corresponding to "peaks". The threshold is a percentage of the set of the extracted points of interest.

III. Experiments

A. Synthetic Data

To generate synthetic data, let us consider a reference signal (line intercept parameter) extracted from a real sequence composed of 800 images. A rupture of the reference signal (x (*i.e.* $x \in \mathcal{G}$)) surely corresponds to a rupture of the position signal. From this reference signal, several position signals are simulated as follows:

Several points belonging to $\mathcal{G}$ are randomly chosen. The percentage of chosen points is called tc. In order to preserve a zero gradient for these points, the latter and their neighborhood are not modified. The other points of the curve are disturbed with trigonometrical functions. Figure 5 illustrates the result obtained by preserving 20 % of points belonging to $\mathcal{G}$.

For these experiments, the latency is zero. The figure 6 illustrates the cost function variation with a fixed value of tc ($tc = 10\%$) for different random position signals. The cost function (see equation 1) at zero may only be a local minimum. This is due to the nature of the cost function: for a given current estimate of the latency l, the cost is the sum of all signal gradients at points $(x + l)$. However, a signal rupture (i.e., a point x in the set $\mathcal{G}$) might not correspond to a rupture for the six position signals. One or more non-zero

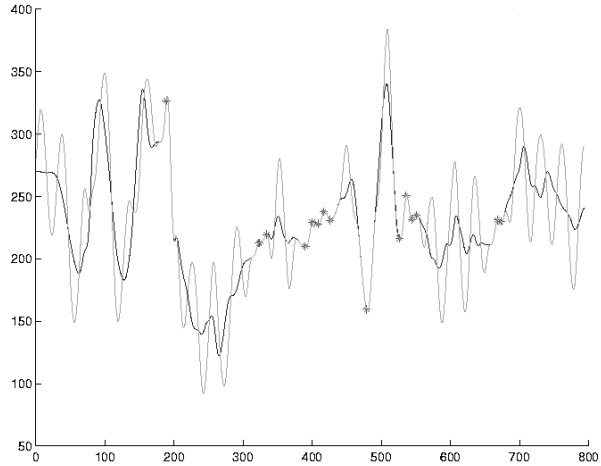


Fig. 5. Synthetic position sensor position (gray) generated using a reference signal (black). Zero gradient points detected are illustrated with stars.

partial derivatives of the signal might cause the presence of local minima.

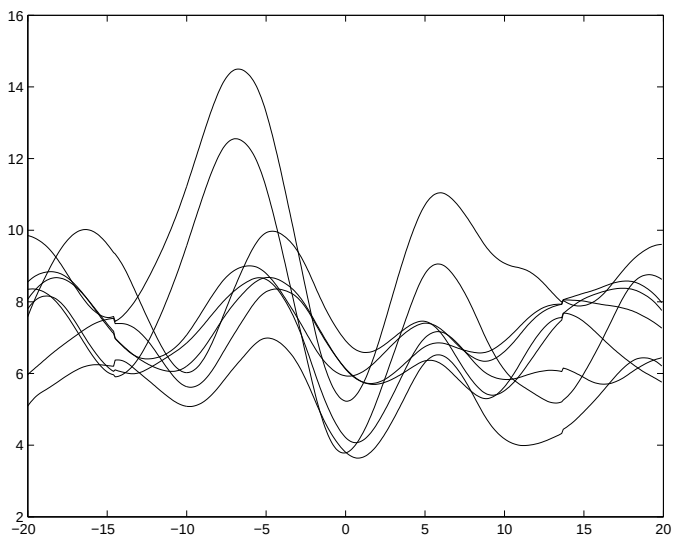


Fig. 6. Top: cost function variation with a fixed tc value ($tc = 10\%$) according to random position signal. Abscissa: latency. Each cost function curve corresponds to one simulation.

The robust formulation allows us to overcome this critical difficulty. Figure 7 illustrates the cost function variation by using M-estimators with tc equal to 10%. A global minimum is now clearly defined, even with low tc.

The latency domain of interest is relatively small: $[-1s, 1s]$. For a maximum acquisition rate set to 25 images per seconds, this interval corresponds to 50 images. A very rough initialization is obtained by a rapid exhaustive search. Accurate latency is then estimated using Powell's optimization algorithm [11].

B. Real Sequences

A Sonosite 180PLUS ultrasound machine (Bothell, USA) and a $C60/5-2$ curved probe (depth $12cm$) were used for the experiments. The probe was tracked using an electromagnetic

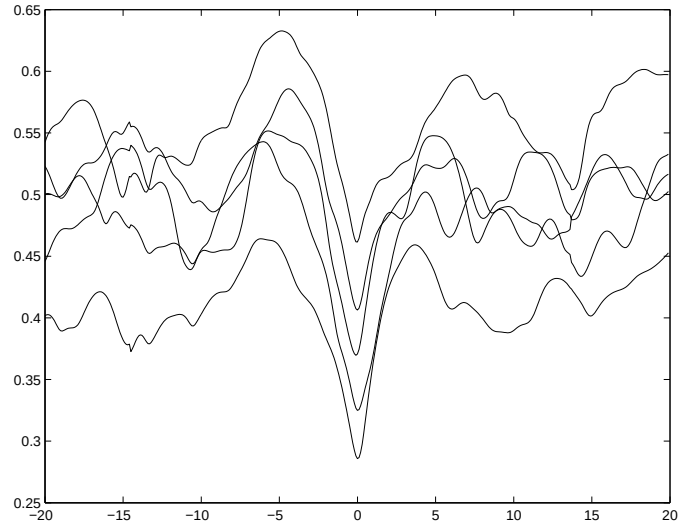


Fig. 7. Cost function evolution w.r.t. latency with a fixed tc value ($tc = 10\%$) using robust estimator according to random position signal. Each cost function curve corresponds to one simulation.

Bird system (Ascension Technology Corporation, Burlington, USA) mounted on it. The ultrasound B-scans were acquired on a Linux PC using a mvDELTA (Matrix Vision, Oppenweiler, Germany) acquisition video card. The StradX software [3] was used to acquire images and 3D position data provided by the Bird system.

We evaluated our algorithm using two types of sequences called "vertical translational" and "rotational". The first sequences are oscillation sequences obtained by maintaining the probe perpendicular to the plane phantom and by applying an oscillatory vertical motion. This is similar to the type of sequence needed for the StradX calibration facility. "Rotational" sequences are obtained by applying an oscillatory rotation (rotation around the axis perpendicular to the ultrasound imaging plane). In both cases, The acquisition rate was $25Hz$ *i.e.*, the time between two images is $40ms$. For each motion, 20 image sequences were acquired.

The figure 8 shows that even using simple "oscillation" sequences, the line extraction stage can be difficult. In particular, the contrast of the plane in the ultrasound image decreases when the motion magnitude increases. The robust line extraction was able to handle this type of sequence.

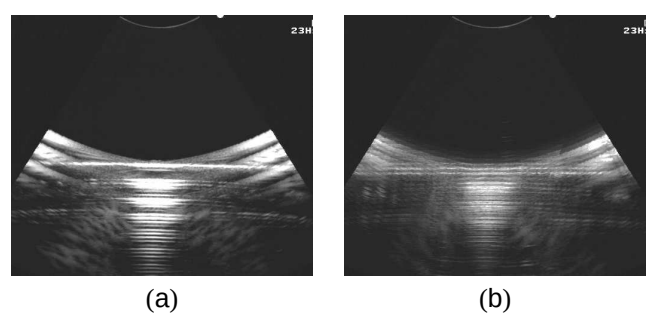


Fig. 8. Ultrasound images of an "oscillation" sequence. (a): first image, (b): image acquired with a vertical probe motion.

The proposed method was compared with the StradX tem-

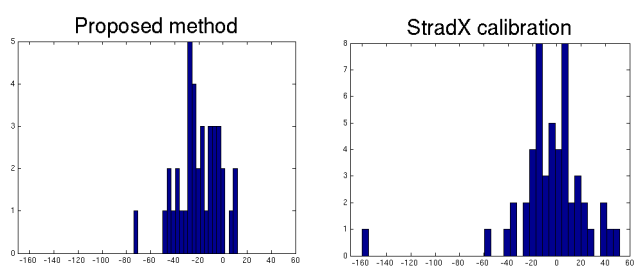
poral calibration method [3]. A temporal calibration of StradX is based on 5 estimates (i.e., 5 image sequences). In this work, 50 image sequences were used and the solution retained for each sequence. Results are presented in table I.

TABLE I

TEMPORAL CALIBRATION (LATENCY IN ms) WITH THE PROPOSED METHOD ("TRANSLATIONAL" AND "ROTATIONAL" SEQUENCES) AND THE STRADX CALIBRATION RESULTS). THE FIRST TWO ROWS PRESENT THE RESULTS FOR A GIVEN MOTION TYPE. THE THIRD ROW PRESENTS THE RESULTS OF ALL SEQUENCES, WHILE THE LAST ROW PRESENTS THE STRADX RESULTS. THE PROPOSED METHOD IS MORE ROBUST AND LEADS TO A LOWER STANDARD DEVIATION OF THE ESTIMATES.

method	mean	median	std. dev.	bounds
"translation" seq.	-17.7	-17.0	20.0	$[-74, 12]$
"rotation" seq.	-22.9	-23.6	15.8	$[-48, 7.6]$
All sequences	-20.4	21.2	18	$[-74, 12]$
StradX	-4.1	-2	30.8	$[-160, 52]$

Estimated latency is small and lower than the image acquisition rate ($40ms$). The latency is negative which means that the ultrasound images were received after the position sensor signals. The proposed method is more robust and leads to a lower standard deviation of the estimates. Figure 9 presents the histograms of the estimates with our method and the StradX method. The estimation bias is lower for the proposed method. The mean estimated latency is slightly different for the two methods. The difference is approximately $16ms$, which is well below the time difference between two images of the sequence ($40ms$). This difference is not related to the way images were acquired, since the StradX software was used in both cases. We suspect that a slight error in line extraction might be the cause of such an error.



(a) (b)

Fig. 9. Histograms of the latency estimates with the proposed method (a) and the StradX method (b). The bias of the estimation is lower with the proposed method. In addition, the true temporal calibration is negative in our case, since images arrive after the position signal (the magnetic position sensor has a much higher rate as the ultrasound machine).

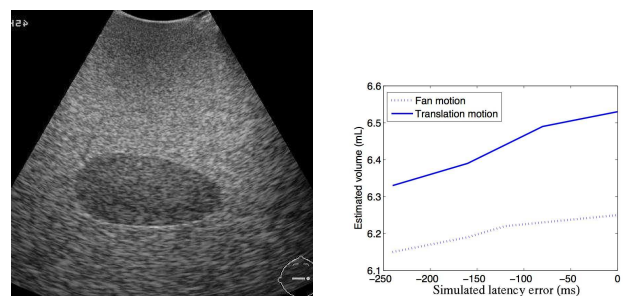
IV. DISCUSSION

A. Validation

The estimated latency of our 3D freehand ultrasound system is lower than $40ms$ (image acquisition rate). It appeared more complex to estimate latency using "spatial calibration" sequences and this for two reasons. Firstly, the probe motion is less constrained for a "spatial calibration" sequence. The probe motion changes may not be clearly detected. The second reason refers to the image acquisition rate. Considering an acquisition rate equal to 25 images per second and an acquisition time of 30s, the size of a "spatial calibration" sequence is approximately equal to $200MB$. This volume of data is significant whereas the acquisition time is short. Then, the image acquisition rate is set to 5 images per second for spatial calibration sequences. This under-sampling obviously downgrades the latency estimation.

B. Latency influence on 3D freehand ultrasound accuracy

The estimated latency of our 3D freehand system is low, but is it negligible? To evaluate latency influence on reconstruction accuracy, a 3D US calibration phantom was scanned (CIRS, model 55. http://www.cirsinc.com). A test object with known volume provided by the manufacturer ($6.6162mm^3$), was reconstructed. Figure 10 shows one B-scan of the phantom. Therefore, we compared the volume given by the manufacturer with the estimated volume for different latency values. For image acquisition of this phantom, the depth was $10cm$, and two different probe motions were used: translational motion and fan-like motion. The image acquisition rate used was 25 images per second. Volume estimation was performed using the manual segmentation tool of StradX. Spatial calibration was done using the method described in [4].



(a) (b)

Fig. 10. Validation with calibration phantom. (a): Image of phantom B-scans, (b): Latency influence on volume estimation (depth $10cm$).

We simulated errors in latency estimation up to $-240ms$. 3D volumes were reconstructed with this temporal mis-calibration. Figure 10 shows the impact on volume estimation. Small latencies (about $-40ms$) do not disturb in a significant way the volume estimation. Such a result could be predicted because during acquisitions, the probe motion was very slow. Norms of the translation vector between two images are approximately equal to $0.2mm$, which is lower than the localization system precision. These experiments show that a small latency about $40ms$ does not introduce a significant error into volume estimation procedure. However, Jacobs reported a latency of $220ms$ [5]. Our study shows that such a large mis-calibration causes an inaccuracy of 5% on the 3D volume estimation and cannot be neglected.

C. Conclusion

We proposed a new approach for temporal calibration for 3D ultrasound systems. The formulation of latency estimation is very general and does not introduce any constraints on probe motion. Latency estimation is stated as a robust minimization issue. Experiments have shown that accurate latency estimation is preferably conducted using ”oscillation” sequences. Experiments have shown that inaccurate latency estimation may induce significant errors in 3D reconstruction.

A future direction is the possibility to perform jointly the temporal and spatial calibration. In effect, this would provide a simple calibration procedure easy to use for 3D freehand ultrasound.

ACKNOWLEDGMENT

The authors would like to thank Brittany region for material support, Wiro Niessen and Marloes Letteboer from the ISI Utrecht for material assistance concerning ultrasound image acquisition, and R. Prager and G. Treece for their help on using StradX software.

REFERENCES

[1] A. Fenster, D. B. Downey, and H. N. Cardinal, “Three-dimensional ultrasound imaging,” *Physics in medicine and biology*, vol. 46, pp. 67–99, 2001.

[2] A. H. Gee, R. W. Prager, G. Treece, and L. Berman, “Engineering a freehand 3d ultrasound system,” *Pattern Recognition Letters*, vol. 24, pp. 757–777, 2003.

[3] R. W. Prager, G. A. H., and B. L., “Stradx: real-time acquisition and visualisation of freehand three-dimensional ultrasound,” *Medical Image Analysis*, vol. 3, no. 2, pp. 129–140, 1999.

[4] F. Rousseau, P. Hellier, and C. Barillot, “Confhusius: a robust and fully automatic calibration method for 3D freehand ultrasound,” *Medical Image Analysis*, vol. 9, pp. 25–38, 2005.

[5] M. C. Jacobs, M. A. Livingston, and A. State, “Managing latency in complex augmented reality systems,” in *Symposium on Interactive 3D Graphics*, 1997, pp. 49–54.

[6] G. M. Treece, A. H. Gee, R. W. Prager, C. J. C. Cash, and L. H. Berman, “High resolution freehand 3d ultrasound,” *Ultrasound in Medicine and Biology*, vol. 9, no. 4, pp. 529–546, 2003.

[7] M. Nakamoto, Y. Sato, K. Nakada, Y. Nakajima, K. Konishi, M. Hashizume, and S. Tamura, “A temporal calibration method for freehand 3d ultrasound system : a preliminary result,” in *CARS*, pp. 1365, 2003.

[8] P. Hough, “Machine analysis of bubble chamber pictures,” in *International Conference on High Energy Accelerators and Instrumentation, CERN*, 1959.

[9] B. Chalmond, *Modeling and inverse problem*. Springer, 2003.

[10] M. Black and A. Rangarajan, “On the unification of line processes, outlier rejection, and robust statistics with applications in early vision,” *Int. J. Computer Vision*, vol. 19, no. 1, pp. 57–91, 1996.

[11] M. Powell, “An efficient method for finding the minimum of a function of several variables without calculating derivatives,” *The Computer Journal*, pp. 155–162, 1964.