

Confhusius: A robust and fully automatic calibration method for 3D freehand ultrasound

François Rousseau, Pierre Hellier ^{*}, Christian Barillot

IRISA, Université Rennes 1-INRIA-CNRS, Campus de Beaulieu, 35042 Rennes cedex, France

Received 10 July 2003; received in revised form 21 April 2004; accepted 9 June 2004

Available online 23 August 2004

Abstract

This paper describes a new robust and fully automatic method for calibration of three-dimensional (3D) freehand ultrasound called Confhusius (CalibratiON for FreeHand UltraSound Imaging USage). 3D Freehand ultrasound consists in mounting a position sensor on a standard probe. The echographic B-scans can be localized in 3D and compounded into a volume. However, especially for quantitative use, this process dramatically requires a calibration procedure that determines its accuracy and usefulness. Calibration aims at determining the transformation (translations, rotations, scaling) between the coordinates system of the echographic images and the coordinate system of the localization system. To calibrate, we acquire images of a phantom whose 3D geometrical properties are known. We propose a robust and fully automatic calibration method based on the Hough transform and robust estimators. Experiments have been done with synthetic and real sequences, and this calibration method is shown to be easy to perform, accurate, automatic and fast enough for clinical use.

© 2004 Elsevier B.V. All rights reserved.

Keywords: 3D freehand ultrasound; Calibration; Geometry; Evaluation

1. Introduction

As it is mostly non-invasive and has a real time capability and a relatively low cost nature, 2D ultrasound is popular. Its major drawback is its weak capability of issuing quantitative accurate morphometric information (Fenster et al., 2001). In fact, conventional ultrasound exams are limited by 2D viewing, and follow-up studies are not easily reproducible. 3D ultrasound imaging overcomes these limitations. In addition, it facilitates extensive investigation and allows accurate measurements of organ volumes.

A 3D examination can be decomposed into three stages: acquisition, reconstruction and visualization. The acquisition stage is crucial because it strongly affects the final results. For quantitative studies, three solutions have been proposed to acquire images: 3D probes, mechanical systems and freehand acquisition systems (manual sweeping). The most straightforward way is to modify classical probe to explore directly the patient in 3D with a 2D array. Although this solution is a promising way for 3D ultrasound, 3D probe resolution is lower than classical 2D probe resolution. 2D phased array of elements is still at a development stage (Deng et al., 2002; Holm, 1997). As a matter of fact, technical problems (interactions between the transducers, cumbersome interfacing, complex electronics, etc.) are currently still a challenging research area. The second option is to automate the probe motion with a mechanical system. It produces sequences of regular echographic images of the patient. Three types of mechanical movements have

^{*} Corresponding author. Tel.: +33 2 99 84 25 23; fax: +33 2 99 84 71 71.

E-mail addresses: froussea@irisa.fr (F. Rousseau), pierre.hellier@irisa.fr (P. Hellier), barillot@irisa.fr (C. Barillot).

URL: <http://www.irisa.fr/vista>.

been used so far: linear, rotational and fan. The 3D reconstruction is then accurate, but these types of systems suffer from poor motion possibilities and large structures are difficult to scan.

Finally, the 3D freehand solution consists in fixing a localization system on the probe, which continuously gives the position and orientation of the probe (Prager et al., 1999; Gee et al., 2003a). This is the type of acquisition system which is considered in this paper. The localization system can be magnetic, optic, acoustic or mechanical (Fenster et al., 2001). The calibration process is independent to the localization system used. Knowing positions and orientations of each B-scan, a 3D image can be reconstructed. Figs. 1 and 2 show two examples of 3D reconstruction and visualization of ultrasound data acquired with a tracked freehand system: an ellipsoid of our evaluation phantom (see Section 5.2) and an arm. The illustration of a scanned arm (Fig. 2) shows a reslice view of the organ which is only available using 3D freehand ultrasound. Freehand systems suffer from possible organ motion during the exam (respiration, heart beat, etc.), a lesser accuracy with respect to systems using a 3D probe and a need of a geometrical calibration. However, a freehand system is cheap, flexi-

ble and close to clinical routine so it can readily be applied in many interventions and surgeries. Contrary to 3D probes and mechanical probes, it also allows the examination of large organs (Gee et al., 2003b).

To avoid geometrical artifacts and distortions in 3D freehand ultrasound, orientations and locations of the acquired 2D scans have to be accurately known which requires accurate calibration. This is particularly important in some contexts where geometrical precision is crucial: image guided interventions for instance.

Calibration is needed to correctly localize a B-scan in 4D ($3D + t$) space and embraces 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 3D space. Whatever the localization system, calibration is crucial because it has a significant impact on the quality of reconstruction. The calibration procedure has to be as easy as possible to perform in clinical context, i.e., automatic, user friendly, easy to operate, robust and fast.

We address in this paper two issues related to 3D freehand ultrasound: design of an automatic, robust, fast and reliable spatial calibration procedure; experi-

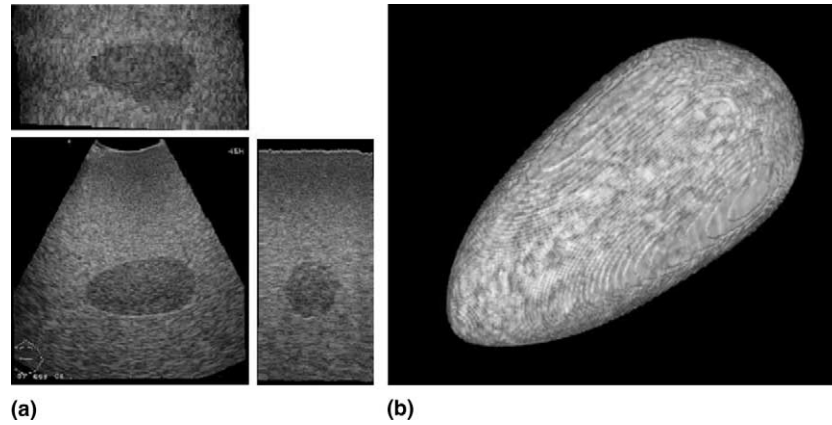


Fig. 1. Ultrasound images of the evaluation phantom: (a) different views of the phantom; (b) 3D visualization of the segmented ellipsoid of the phantom.

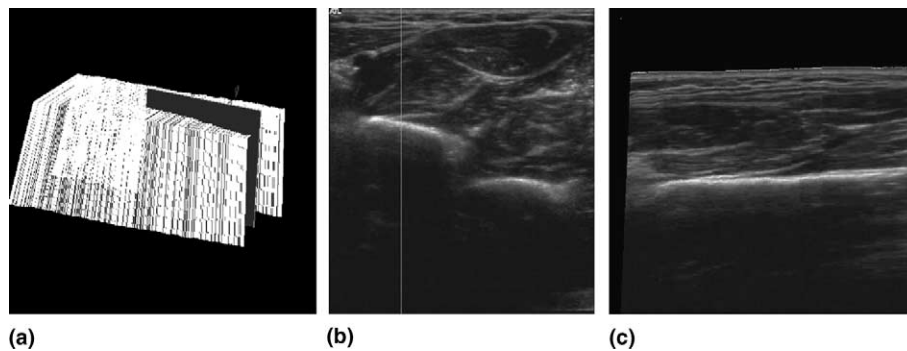


Fig. 2. Arm ultrasound sequence: (a) ultrasound sequence and reslice image in 3D space; (b) classical B-scan; (c), reslice image using StradX software (Prager et al., 1999).

ments with synthetic and real sequences and a 3D ultrasound calibration phantom. The paper is organized as follows. Section 2 presents the spatial calibration problem and related work is described in Section 3. The proposed calibration method is described in Section 4. Section 5 presents the experiments and the results.

2. Spatial calibration formulation

The spatial calibration problem can be formulated in the following way (Prager et al., 1998):

$$\mathbf{x}_c = T_c T_t T_r \mathbf{x}_r, \quad \text{with } \mathbf{x}_r = \begin{pmatrix} s_x u \\ s_y v \\ 0 \\ 1 \end{pmatrix}, \quad (1)$$

where T_r denotes the rigid transformation from B-scan to receiver coordinates, T_t is the rigid transformation from receiver to transmitter coordinates, T_c stands for the rigid transformation from transmitter to phantom coordinates, u and v are the image pixel coordinates, s_x and s_y denote scaling factors or pixel spatial resolution (see Fig. 3).

Performing the calibration amounts to estimating the matrix T_r (three translations and three rotations) and the scaling coefficients s_x and s_y . The calibration is generally carried out in the following way: after scanning a phantom whose 3D geometrical properties are known, the calibration is based on this geometry being recovered in the sequence of ultrasound images.

T_r is a rigid-body transformation:

$$T_r(t_x, t_y, t_z, \alpha, \beta, \gamma) = \begin{pmatrix} R(\alpha, \beta, \gamma) & \mathbf{t}(t_x, t_y, t_z) \\ 0 & 1 \end{pmatrix}, \quad (2)$$

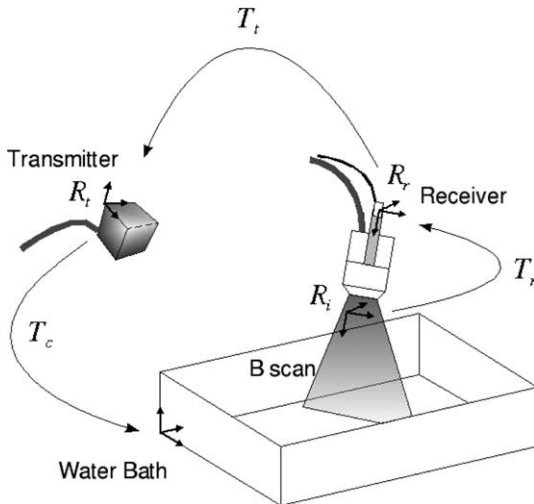


Fig. 3. Coordinate systems: R_i (image), R_r (receiver), R_t (transmitter), R_c (reconstruction volume, here the water bath).

where $\mathbf{t}(t_x, t_y, t_z)$ is a translation vector, (α, β, γ) the rotation angles around axis (z, y, x) and $R(\alpha, \beta, \gamma)$ is a rotation matrix (see Appendix A).

3. Related work

Spatial calibration methods can be classified in three groups according to the phantom used: wire phantom, single-wall phantom, multimodality registration phantom.

3.1. Wire phantom

3.1.1. Cross-wire

The most popular point based calibration method (Detmer et al., 1994; Barry et al., 1997; Meairs et al., 2000) is the cross-wire method proposed by Detmer et al.: the intersection of two wires is scanned. Calibration parameters are estimated by minimizing the difference between the recorded positions of the intersection point $\mathbf{P} = (P_x, P_y, P_z)$ and its mean position $\bar{\mathbf{P}} = (\bar{P}_x, \bar{P}_y, \bar{P}_z)$. The total squared error L which has to be minimized is

$$L = [\mathbf{P} - \bar{\mathbf{P}}]^T [\mathbf{P} - \bar{\mathbf{P}}],$$

with

$$\mathbf{P} - \bar{\mathbf{P}} = (P_{x1} - \bar{P}_x, P_{y1} - \bar{P}_y, P_{z1} - \bar{P}_z, P_{x2} - \bar{P}_x, \dots, P_{zN} - \bar{P}_z)^T.$$

The cross-wire phantom is easy to design and the calibration estimation method is quite simple. However, for each B-scan, the intersection point has to be located manually in the image and scanning the point of interest P can be a difficult task. Carr (1996) used a three-wire phantom which consists in three wires mounted in orthogonal directions. Calibration parameters are estimated using this property of orthogonality. It is easier to scan but the design of the phantom requires more attention.

3.1.2. Fiducial based

In (Sato et al., 1998), a specially shaped board is designed for the calibration. It is located under water and three tips are used as control points for the calibration procedure. The tips of the board are localized in the acquired ultrasonic image and their 2D coordinates are specified manually. The 3D positions of these three tips are measured using a Optotrack pen probe. Using the three control points, calibration parameters are estimated.

In (Bouchet et al., 2001), 39 nylon wires forming 13 Z-shaped patterns at 13 different depths are used and are precisely located relative to the base of the phantom with an infrared tracking system. The image position of

each wire is segmented by thresholding. Calibration parameters are estimated by minimizing a least square equation.

In (Pagoulatos et al., 2001), a special calibration phantom is designed and consists of 18 hemispherical divots, nylon strings and a set of 30 fiducials, each with the shape of the letter “N”, called N-fiducials. The N-fiducials are grouped in five planes, each containing six N-fiducials. After registering the phantom with the transmitter coordinate system using the divots in the phantom and a stylus, an US image is acquired, and the intersections of the N-fiducials are identified manually.

Fiducial based methods are quite simple, but each phantom is particular and quite complex. Moreover, the phantoms have to be carefully designed and registered in the transmitter coordinate system; this step can be tedious.

3.1.3. Optically tracked pointer

In (Muratore, 2001), no phantom is needed in the beam-calibration process because the method relies on the spatial relationship that exists between an optically tracked pointer and a similarly tracked US transducer. The pointer tip is placed into the US beam. The US beam is modeled as a 2D plane in 3D space. Calibration parameters are estimated by registering the US image points to the physical points given by a pointer. The image collection and the identification of the pixel coordinates of the pointer in B-scans are time-consuming processes. The entire calibration procedure takes approximately 60 min.

3.2. Single-wall phantom

The idea of using a wall of a water bath to calibrate a 3D ultrasound system is appealing: the single-wall phantom is very simple to build and easy to use. Moreover, the water bath plane appears to be a straight line in a B-scan. An extracted line defines two points on the phantom’s plane. Each pixel on the line should satisfy (Prager et al., 1998)

$$\begin{pmatrix} x \\ y \\ 0 \\ 1 \end{pmatrix} = T_c T_t T_r \begin{pmatrix} s_x u \\ s_y v \\ 0 \\ 1 \end{pmatrix}. \quad (3)$$

The zero component of this equation gives one equation to determine the calibration parameters. So, given a line defined by two pixels, each B-scan gives two equations. Thus, using n B-scans, the over-determined set of $2n$ equations can be solved.

Prager et al. extract semi-automatically a straight line using gradient information in each B-scan (Prager et al., 1998). The line detection step has to be manually verified

for each B-scan and some parameters have to be tuned for extracting good lines (gradient threshold, RANSAC acceptance proportion, etc.). No 3D spatial coherence is checked between all the extracted lines and it might be a problem with noisy sequences to reach an acceptable precision for calibration parameters.

The method developed by Prager et al. (1998) is efficient, freely available and can be considered as a reference in this domain. However, it requires user intervention and care about the experimental condition. We have experienced that the degree of user-intervention must be as low as possible for a clinical context. Finally, they have developed a special phantom (Prager et al., 1998) to overcome the problem linked to the beam thickness, but this new phantom has to be built and is not so easy to manipulate for clinical routine.

3.3. Image registration approach

Blackall et al. (2000) recently proposed an image registration approach to calibrate freehand 3D ultrasound. The idea is that a consistent calibration gives an optimal similarity measure between the ultrasonic images of the phantom and a 3D model of this one (CT or MR volume).

$$\hat{T}_r = \arg \max_T \{I(A, T(B))\}, \quad (4)$$

$I(A, T(B))$ is the similarity function between the set of ultrasound images B and the digital model A .

The MRI volume and the US volume have to be expressed in the same 3D coordinate system. So, it is necessary to know the coordinates of the landmarks in the space of reconstruction. This can be done with a high degree of accuracy using an optical tracker but this step is tedious and time-consuming. Then, using the normalized mutual information as similarity measure, registration between the MRI volume and the US volume permits to estimate the calibration matrix.

3.4. Discussion on literature and motivations of our work

In (Prager et al., 1998), a comparison between cross-wire method three-wire method and single-wall method is proposed. The cross-wire and the three-wire techniques suffer from relatively poor repeatability and extremely long calibration times. It has been shown (Prager et al., 1998) that using a flat plane produces lines that can be reliably and automatically detected. In (Blackall et al., 2000), the final results of precision and reproductibility are comparable with the method using cross wire phantom, but with a reduced acquisition time (2 min). Although this method is interesting, setup is complex (MRI acquisition, reliability of registration between US and MRI, specific phantom) and calibration parameters have to be initialized quite close to the solu-

Table 1
Advantages and drawbacks of existent calibration methods

Calibration method	Cost	Ease to build	Ease to use	Quickness	Automatic
Cross-wire	+	+	–	–	No
Fiducial based	+	–	+	–	No
Optically tracked pointer	–	+	–	–	No
Single-wall phantom	+	+	+	–	No
Image registration approach	–	+	–	–	No

tion. Finally, although other interesting methods have been proposed, they were not compared.

Table 1 summarizes advantages and drawbacks of existent calibration methods. We take into account several criteria to characterize calibration method:

- the cost of the needed material to do calibration (phantom, optical tracker, etc.),
- the simplicity of the used phantom, the ease to build it,
- the ease to do the calibration (degree of expertise of user),
- the quickness of the calibration process,
- the need for external intervention (user interaction, automaticity).

We think there is a clear need for a cheap, robust, fast and fully automatic procedure. In order to use 3D ultrasound in routine applications, we propose a calibration technique that complies with these requirements. This is particularly needed when the attachment between the probe and the localization system cannot be stable over time.

4. Method

4.1. Principle

With the aim of carrying out a simple, fast and completely automatic calibration, we have chosen to use a plane phantom. Fig. 4 shows this phantom which consists of a plexiglass plate in a water bath. This type of phantom is easy to build and to use. Moreover, it provides in each image a strong, straight line which can be automatically detected with high accuracy. This kind of phantom does not need any registration with optical tracker in the reconstruction coordinate system.

First, a point set in each image coordinate system is extracted (see Section 4.2 for the extraction stage explanation) and this point set should correspond to a plane. Second, given a plane model in the reconstruction coordinate system, we estimate the optimal rigid transformation that aligns the plane and the point set. This makes it possible to estimate both plane parameters and calibra-



Fig. 4. Plane phantom used to calibrate the 3D freehand ultrasound system.

tion parameters (see Section 4.3). Lastly, we use an automatic outlier rejection process for the extraction and optimization steps in order to obtain an automatic and robust method.

4.2. Extraction

In each B-scan, we extract points of interest belonging to the line produced by the water bath (see Fig. 8). The entire extraction process is synthetized in Fig. 5.

The highest gradient and highest luminance points are retained by thresholding. The threshold is fixed (in percentage) compared to the maximum in intensity and gradient in the first image. These two features are needed because, as it can be seen in Fig. 8(b), either intensity or gradient information alone would not be sufficient for a reliable extraction.

Since the points of interest extraction is done by thresholding, this step is sensitive to outliers. At this stage, a 3D outliers rejection might remove entire B-scan sequences. This would be damageable since a correct calibration relies on a complete set of probe motion (see Prager et al., 1998). To obtain a robust algorithm, the consistency of the extraction is checked successively using 2D information (Hough transform) and 3D information (spatio-temporal continuity of the point set).

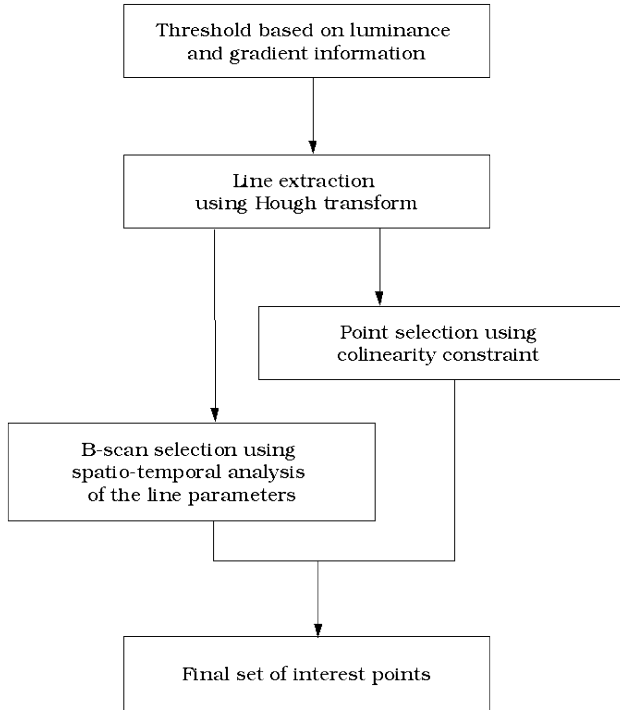


Fig. 5. Extraction procedure.

4.2.1. 2D consistency: the Hough transform

To reject outliers in each image (see Fig. 8(b)), the Hough transform is used (Hough, 1959). The Hough transform can isolate features of a particular shape within an image. Because it requires that the desired features be specified in some parametric form, the classical Hough transform is most commonly used for the detection of regular curves such as lines, circles, ellipses, etc. A convenient equation for describing a set of lines uses parametric or normal notion: $x \cos \theta + y \sin \theta = r$, where r is the length of a normal from the origin to this line and θ is the orientation of r with respect to the X -axis in the B-scan. For any point (x, y) on this line, r and θ are constant. When viewed in Hough parameter space (r, θ) , points which are colinear in the cartesian image space become readily apparent as they yield curves which intersect at a common (r, θ) point. The transform is implemented by quantizing the Hough parameter space into finite intervals or accumulator cells. Each

(x_i, y_i) is transformed into a discretized (r, θ) curve and the accumulator cells which lie along this curve are incremented. Peaks in the accumulator array represent strong evidence that a corresponding straight line exists in the image.

The Hough transform is used to extract in each B-scan a line whose equation is: $ax + by + c = 0$. Given the equation of the line and the point set, we have to reject possible outliers. Using the Euclidian distance as selection criterion, only the closest points to the line are retained using a fixed threshold.

4.2.2. 3D consistency: spatio-temporal continuity of the point set

To ensure coherence in all the ultrasound sequence, we reject image where the line extraction is not correct. Continuity of probe motion should lead to a smooth variation of parameters a , b and c . Fig. 6 plots the evolution of a , b and c for the US sequence. We use a parametrization of the line using a couple of parameters $(a/b, c/b)$. In order to reject incorrect images which contain bad detected lines, the evolution of line parameters are analyzed. First, B-spline approximation of the evolution of line parameters is performed. For details concerning B-spline approximation of data set, see (Chalmond, 2003). Then, the differences between these approximations and the line parameter evolutions which we will call “difference signals X_1 and X_2 ” are considered. A hypothesis test based on these difference signals is performed to reject outliers (i.e. incorrect images). The mean of the “difference signals X_i ” is supposed to be equal to zero. We assume that observations X_i are independent and follow the same law $\mathcal{N}(0, \sigma^2)$ where σ is unknown. As the mean is known, $T = (1/n) \sum_{i=1}^n (X_i - m)^2$ is the best estimator of σ^2 . The type 1 risk α is chosen equal to 0.05. Thus, if $|X_i| / \sigma > 1.96$, the observation X_i is rejected (Lehmann, 1986), see Fig. 7.

4.2.3. Comparison with StradX (Prager et al., 1998)

Prager et al. have developed a calibration method (included in the StradX software) using a plane phantom. As explained above, a plane phantom is easy to use and provides features clearly detectable in each B-scan.

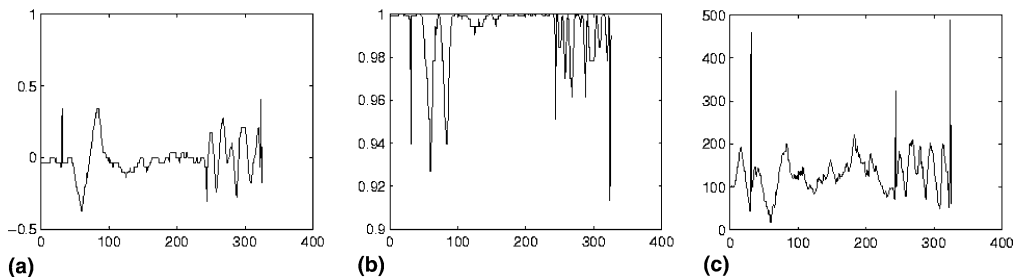


Fig. 6. Evolution of the line parameters (a)–(c) for a classical calibration sequence.

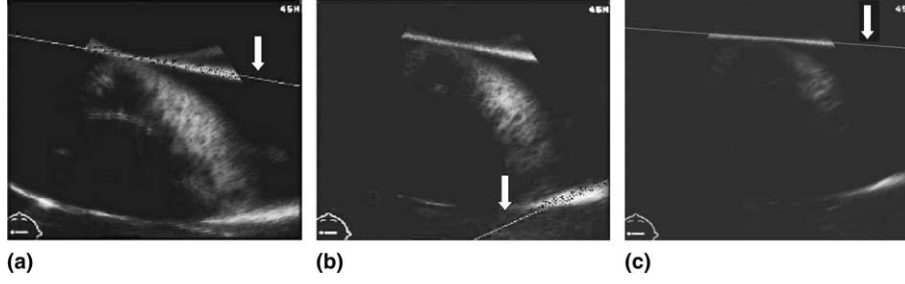


Fig. 7. Line detection with the Hough transform for three successive images: (a) and (c) good detection; (b) bad detection, the line is rejected.

First, the line detection process developed by Prager et al. uses only 2D information and is based on gradient information. A point set of interest is extracted using n vertical lines which are sampled at regular intervals in the ultrasound image. The intensities along each line are smoothed using a Gaussian filter, and the variance of the kernel has to be tuned. Threshold of gradient information has to be chosen by the user. The extraction process proposed by Prager et al. is not totally automatic and intervention of the user is required. In our case, gradient and luminance are combined to extract interest points and thresholds are fixed by using the first image of the sequence as a reference. Moreover, we consider the entire image to extract the point set.

Random sample consensus (RANSAC) (Fishler and Bolles, 1981) is used to fit a line to the data. A point contributes to the consensus set if it is no more than a distance d above or below the line. The user has to tune the RANSAC accept threshold and the parameter d . We propose to fit a line to the data to use the Hough transform which is well-known for its robustness. Moreover, the results of the extraction process is not sensitive to the variations of Hough parameters (quantification of the parameter space) and only one parameter (the Euclidian distance used as selection criterion) has to be tuned.

Finally, StradX requires user intervention to check line detection results over all the sequence, and to tune line detection parameters for the 2D outlier rejection process. In our case, combining 2D and 3D information and using spatio-temporal continuity of the point set makes the method fully automatic which is essential for clinical use.

4.3. Formulation

In clinical routine conditions, the position of the plane in the reconstruction volume is unknown. Since it is desirable to make as few assumptions as possible, the parameters of the plane have to be estimated. Therefore, Eq. (1) becomes: $\mathbf{x}_t = T_t T_r \mathbf{x}_r$.

Calibration parameters and plane parameters are estimated by minimizing the Euclidian distance between the plane and the points of interest. We propose here two different formulations to do so:

1. A distance criterion expressed in 3D coordinate system:

$$\hat{T}_r = \arg \min_T \left\{ \frac{1}{2} \sum_{i=1}^N d_{3D}(\pi, M_i)^2 \right\}, \quad (5)$$

where N is the cardinal of the point set, $d_{3D}(\pi, M)$ is the 3D Euclidian distance between the plane π and the point of interest M_i in 3D space.

2. A 2D distance expressed in each B-scan:

$$\hat{T}_r = \arg \min_T \left\{ \frac{1}{2} \sum_{i=1}^N d_{2D}(P(\pi), m_i)^2 \right\}, \quad (6)$$

where $d_{2D}(P(\pi), m)$ is the 2D Euclidian distance between the plane π projected in the image and the point of interest m_i in US images.

In both cases, this complicated nonlinear cost function has to be optimized over a nonlinear parameter space. In both cases, the cost function can be readily derived with respect to all parameters.

The first formulation using 3D Euclidian distance is quite intuitive because the point set has to be aligned with the wall of the water bath (a 3D object). In visual reconstruction domain where 3D structure and viewing parameters (camera pose or/and calibration) have to be jointly estimated, the cost function is classically defined in images using a basic image projection model (Triggs et al., 1999). The reason is that the 3D coordinate frame is itself uncertain, as it can only be located relative to uncertain reconstructed features or cameras. In (Laveau, 1996), the two formulations have been used in the context of computational projective geometry and experiments have shown that there are only few differences. Concerning the calibration problem, there is no reason to favor one method. The two formulations will be experimentally compared in Section 5.

4.4. Initialization

In the first image, the probe is broadly perpendicular to the water bath. It can be assumed in a first approximation that the plane π is orthogonal to the image plane (this assumption is relaxed afterward). Plane parameters

can be initialized using 3D coordinates of the first image. Calibration parameters should be initialized with respect to the geometry of the clip between the probe and the sensor.

4.5. Optimization

Using Levenberg–Marquardt algorithm (Moré, 1978), the minimization process consists in estimating all the calibration parameters and coefficients plane.

To shorten the computation time, we propose a hierarchical algorithm for both 3D and 2D formulations. The complexity of the algorithm only depends on the number of points of interest used to minimize f . Classically in computer vision, to shorten computation time, a multiresolution algorithm is used (Burt, 1984). In a similar way, we can divide the optimization procedure into different stages. For instance, with a fixed number N of interest points, the first optimization part is performed with a third of N points. This result is used as an initialization for an optimization using two third of N points. Then the second part with two thirds of N points, and finally, with all extracted points. This hierarchical decomposition permits us to split the algorithm complexity and in this way, to shorten computation time. Experiments on synthetic sequences have been conducted to determine the optimal decomposition.

4.6. Robustness

Our algorithm uses a least square minimization method. However, it might be very sensitive to outliers due to speckle noise in US image. We have checked for the 2D and 3D consistency of the extracted point set (see Section 4.2). Nevertheless, the set of interest points has been retained by thresholding what is sensitive to outliers. Since calibration parameters are unknown yet, a 3D outlier rejection method cannot be used prior to the optimization process (see Section 4.2). Therefore, we use a outlier rejection process during the optimization stage.

We use least trimmed squares estimator, LTS (Rousseeuw and Leroy, 1987). The LTS estimator \mathcal{T} minimizes a criterion \mathcal{C} with $r_1^2 \leq r_2^2 \leq \dots \leq r_N^2$ the ordonated remainders, and h the number of points used for the estimation,

$$\hat{T} = \arg \min_T \sum_{i=1}^h r_i^2, \quad \text{with } \frac{N}{2} \leq h \leq N. \quad (7)$$

The breaking point P_{breakage} and effectiveness \mathcal{E} are two fundamental concepts which characterize estimators:

- $P_{\text{breakage}} = p_{\text{max}}/n$, where p_{max} is the maximal number of measures which the value cannot change without \mathcal{T} norm becomes arbitrarily large, n is the number of points.

- $\mathcal{E} = \sigma_{\text{min}}^2 / \sigma^2$ where σ_{min}^2 is the best estimator variance and σ^2 is the variance of \mathcal{T} .

The breaking point of LTS estimator is equal to 0.5 and \mathcal{E} is close to 1. For each point, computed remainder is equal to the distance between the plane and the point. This allows us to introduce 3D information in the 2D rejection algorithm.

We use LTS estimator because there is only one parameter to tune (the percentage of outliers in each B-scan which is restrictive) contrary to the M -estimators where rejection parameters have to be carefully tuned.

5. Calibration experiments

Our algorithm has been tested with synthetic and real sequences. Synthetic images are interesting because the calibration parameters to be recovered are exactly known. Thus, we can evaluate the precision, the convergence domain and the repeatability of the calibration algorithm. Moreover, it is possible to test the intrinsic precision of our method, because with synthetic sequences, many types of errors are eliminated (localization errors of each B-scan, errors due to beam thickness, etc.). In this section, we will evaluate the hierarchical optimization scheme, the convergence domain, the accuracy and the reproducibility of the method using 2D and 3D formulations. The proposed method is compared with StradX software (Prager et al., 1998) which is a reference in the calibration domain.

5.1. Synthetic sequences

5.1.1. Experimental framework: images and evaluation criteria

The method was tested with three synthetic sequences of 61 images with different calibration parameter sets. In each case, the size of the image is 256×256 (pixels). Each sequence contains 121 images: one reference image and 20 images per motion (three translations and three rotations). Ultrasound images are classically corrupted by the presence of speckle noise. To test the algorithm with more realistic sequences, we add speckle noise to images. Fully developed speckle can be modeled with the Rayleigh law (Dutt, 1995). The Rayleigh probability density function is defined by the following equation:

$$f(x) = \frac{x}{\sigma^2} e^{-x^2/2\sigma^2}, \quad x \geq 0. \quad (8)$$

Speckle is generally modeled as a multiplicative noise: $I_2 = nI_1$, where I_1 is the original image, I_2 is the corrupted image and n is the speckle pattern. As speckle noise is a multiplicative noise, we add noise to a gray background of the image where σ is equal to 1. Around the plane,

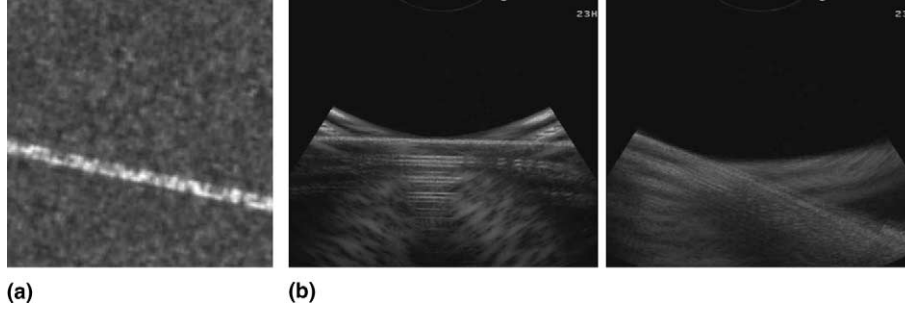


Fig. 8. Representative images of US sequences: (a) added speckle noise synthetic image; (b) real images of the water bath.

noise is increased ($\sigma=10$). Fig. 8(a) shows a B-scan extracted from synthetic noisy sequence.

The initialization stage of the algorithm is an important part of the method that we have to evaluate. For each parameter, a initialization domain has been defined with respect to known optimal parameters. For each test, the algorithm was tested with 40 starting estimates, with combinations of ± 40 mm and $\pm 40^\circ$ for the rigid transformation parameters, and $\pm 50\%$ for the scaling factors.

5.1.2. Evaluation criteria

To evaluate the quality of the results for the synthetic sequences, mean and standard deviation of e_t, e_r, e_s are calculated. e_t, e_r, e_s are defined as follows:

$$e_t = \|dt\|_2, \quad \text{with } dt = \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} - \begin{pmatrix} \hat{t}_x \\ \hat{t}_y \\ \hat{t}_z \end{pmatrix},$$

$$e_r = \|dR\|_2 = \sqrt{\text{trace}(dR^t dR)}, \quad \text{with } dR = R - \hat{R},$$

$$e_s = \|ds\|_2, \quad \text{with } ds = \begin{pmatrix} s_x \\ s_y \end{pmatrix} - \begin{pmatrix} \hat{s}_x \\ \hat{s}_y \end{pmatrix},$$

(t_x, t_y, t_z) , R and (s_x, s_y) are the true calibration parameters. e_t is the translation error with respect to (t_x, t_y, t_z) , $(\hat{t}_x, \hat{t}_y, \hat{t}_z)$ are the estimated translation parameters. e_r is the rotation matrix error with respect to R , \hat{R} is the estimated rotation matrix. e_s is the scale factors error with respect to (s_x, s_y) , (\hat{s}_x, \hat{s}_y) are the estimated scale factors. e_t is expressed in mm, e_r in degree and e_s in mm/pixel.

Mean error indicates the convergence of the algorithm while the standard deviation indicates the reliability of the results. For each experiment, the method has been tested with 20 sets of initialization parameters.

5.1.3. 3D localization of a point

In order to quantify 3D reconstruction errors due to a miscalibration estimation, experiments about the 3D point localization accuracy of a freehand system have been performed. A single point is considered in 3D

space. 50 positions of the probe are simulated and the difference between the real position of the point and the position estimated with the 3D freehand system is computed. Fig. 9 illustrates this experiment.

The noise of calibration parameters is modeled with the normal law. Table 2 shows the four kinds of simulated disruptions. The simulation results are given in Table 3. These results show that the calibration parameter estimation has to be the most accurate as possible. Indeed, in the third simulation ($\sigma_{\text{translation}}=2$ mm, $\sigma_{\text{rotation}}=2^\circ$, $\sigma_{\text{scale}}=10\%$), the mean of the 3D localization error is equal to 3.270 mm. The required accuracy of a 3D freehand ultrasound system depends on the medical application. However, this experiment shows the importance of the calibration procedure for a quantitative use of 3D freehand ultrasound.

5.1.4. Hierarchy

Figs. 10 and 11 show the evolution of e_t , e_r and e_s according to the applied sub-sampling using 2D and 3D minimization criterion for one synthetic sequence. The plain curve represents the mean error, the dashed lines represent the mean plus or minus one standard deviation. Concerning 3D criterion, we can note that using the sub-sampling strategy, e_t , e_r and e_s definitely decrease. Using 2D criterion, e_t is decreasing with the sub-sampling, e_r and e_s are stable and small. It seems valid to divide the optimization into three stages.

5.1.5. Initialization domain

The initialization of the algorithm is an important part of the method that we must evaluate. For each parameter, the initialization domain was defined according to the parameters to estimate. To assess the algorithm stability with regard to the initialization stage, six initialization domains have been tested (see Table 4).

Fig. 12 (resp. Fig. 13) shows for one synthetic sequence the evolution of e_t , e_r and e_s according to the tested initialization domain using 2D (resp. 3D) criterion. One uses the same representation as for Fig. 10. The method using 3D minimization criterion is clearly more stable than the method using 2D criterion. We can note that even when the initialization domain is

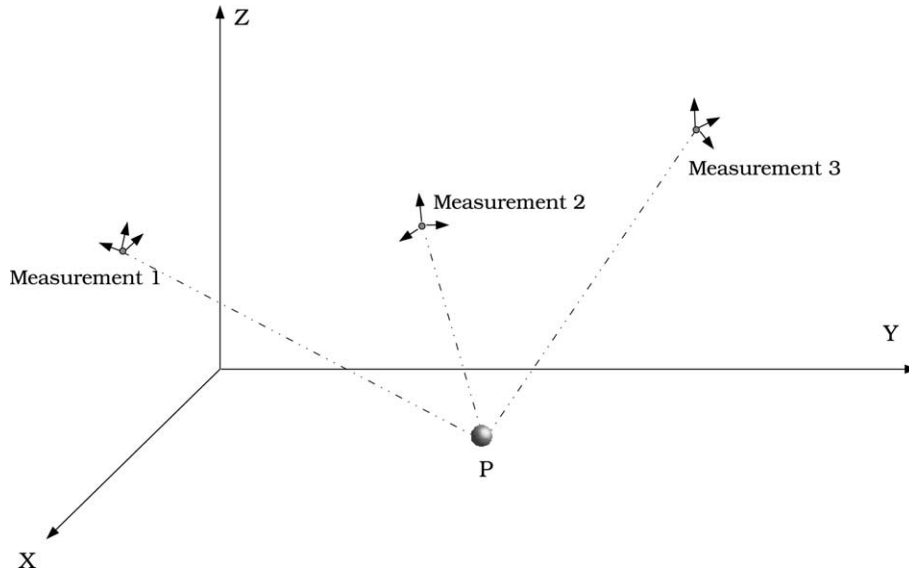


Fig. 9. Diagram illustrating multiple location measurements of a single point P . 3D localization accuracy is estimated by computing the mean difference between estimated positions and real positions of the point.

Table 2

Samples are generated using a Gaussian law

Test domains	Translation (mm)	Rotation ($^{\circ}$)	Scale (%)
1	0.5	0.5	5
2	1	1	10
3	2	2	10
4	3	3	15

For each test domain, the standard deviation values of translation, rotation and scale are shown.

Table 3

3D localization errors (distance between estimated position and real position) for each test domain defined in Table 2

	Exp. 1	Exp. 2	Exp. 3	Exp. 4
Mean (mm)	0.824	1.626	3.270	4.937
Standard deviation (mm)	0.346	0.705	1.400	2.090
Maximum (mm)	2.252	4.102	8.539	13.216
Minimum (mm)	0.060	0.132	0.146	0.411

vast, the algorithm behaves well (whatever the minimization criterion used).

5.1.6. Accuracy and reproductibility

For each experiment, the algorithm has been tested with 20 different initialization domains (± 40 mm for

translation, $\pm 20^{\circ}$ for rotations and $\pm 20\%$ for scale factors). Calibration results are summarized in Table 5.

These results indicate that even with noisy sequences and a large initialization domain, the calibration parameters estimation is accurate.

We have tested the three synthetic noisy sequences with the StradX software, developed by Prager et al.

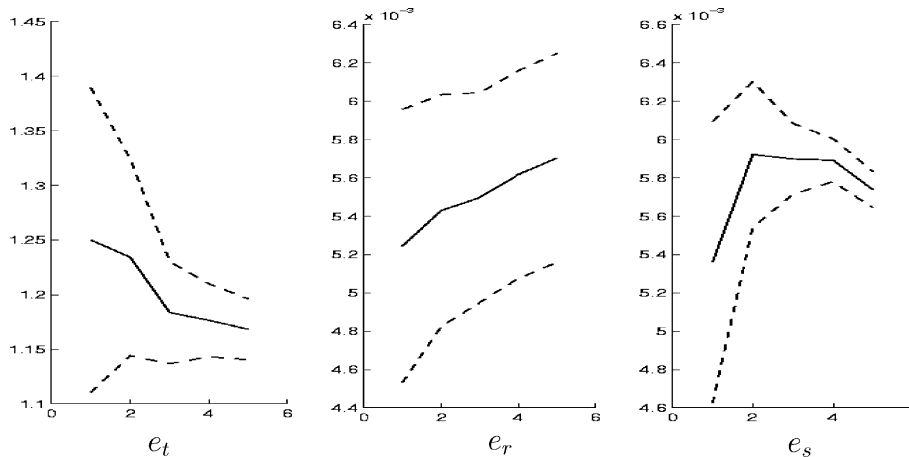


Fig. 10. Evolution of e_t , e_r and e_s according to the applied sub-sampling using 2D criterion. Abscissa: number of step during the optimization, ordinate: mean error; e_t is expressed in mm, e_r in degree and e_s in mm/pixel.

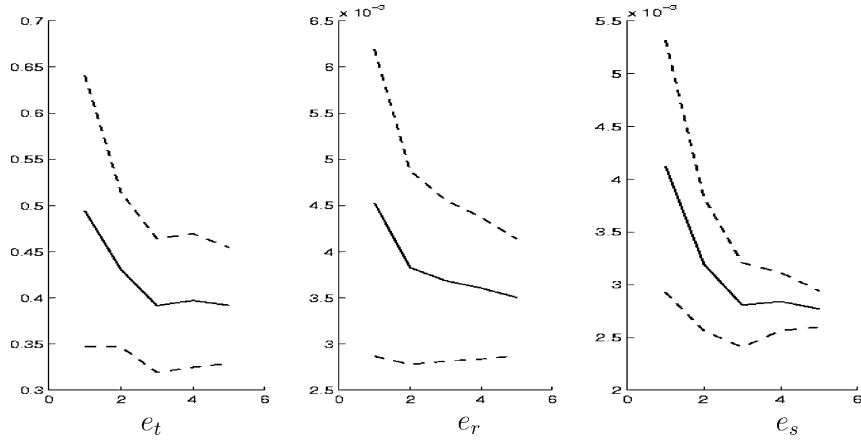


Fig. 11. Evolution of e_t , e_r and e_s according to the applied sub-sampling using 3D criterion. Abscissa: number of step during the optimization, ordinate: mean error; e_t is expressed in mm, e_r in degree and e_s in mm/pixel.

Table 4
Initialization domains tested on synthetic sequences

Domain	Translation (mm)	Rotation (°)	Scale (%)
1	5	5	10
2	10	10	20
3	20	20	30
4	30	30	40
5	40	40	50
6	80	40	50

(1998) at Cambridge University. StradX supplies equivalent results in the other methods of calibration existing. The results are summarized in Table 5. With the presence of fully developed speckle in each sequence, line detection parameters in StradX have to be tuned, otherwise StradX algorithm could fail to converge. Thus, StradX has been tested with different parameter sets tuned to the best of our expertise.

Considering the three synthetic sequences, we conclude that the proposed method seems slightly more suc-

cessful than StradX. We now present results on real data.

5.2. Real sequences

5.2.1. Acquisition

Real sequences have been acquired with Aloka SSD-5000 ultrasound scanner and an optical localization system (Polaris, Northern Digital) attached to the probe. This type of device provides a flexible and very accurate solution. Thus, tracking system errors are minimized (Gee et al., 2003a). Such system has already been successfully widely used for 3D ultrasound (e.g. Sato et al., 1998; Blackall et al., 2000; Bouchet et al., 2001; Muratore, 2001). The probe used in this study was a 7.5 MHz cranial probe.

5.2.2. Calibration evaluation

We evaluate calibration accuracy using a 3D calibration phantom and calibration reproducibility.

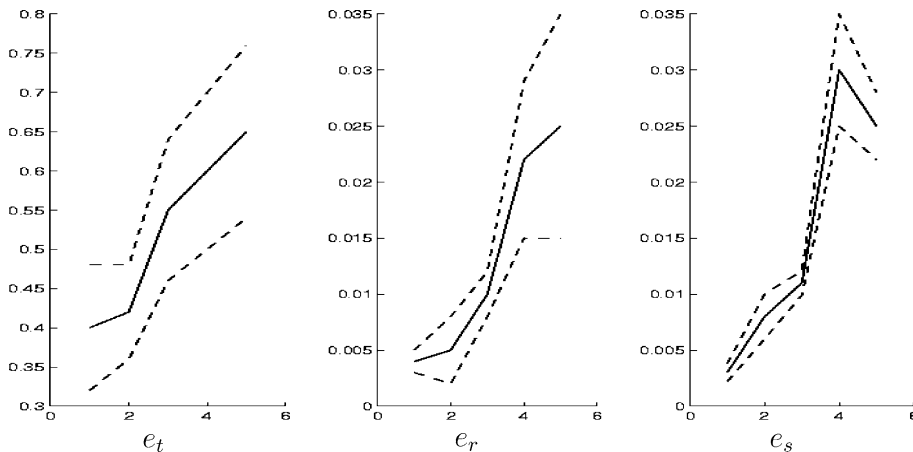


Fig. 12. Evolution of e_t , e_r and e_s according to the tested initialization domain using 2D criterion. Abscissa: tested initialization domain (see Table 4), ordinate: mean error; e_t is expressed in mm, e_r in degree and e_s in mm/pixel.

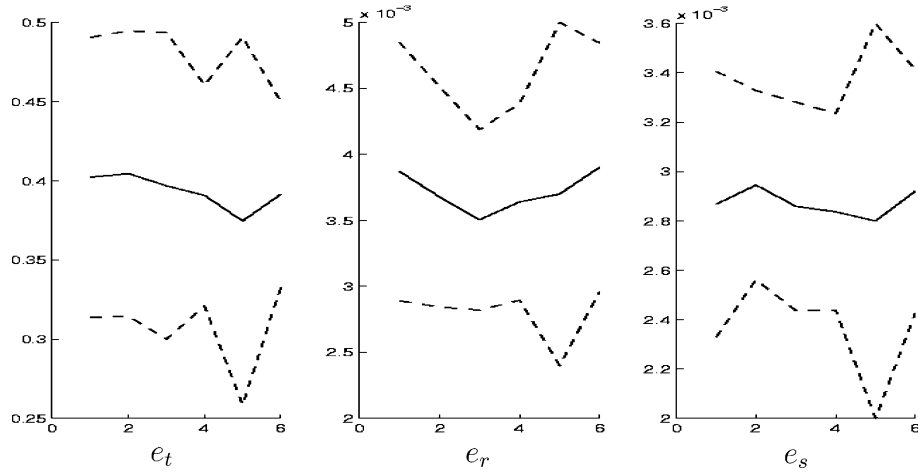


Fig. 13. Evolution of e_t , e_r and e_s according to the tested initialization domain using 3D criterion. Abscissa: tested initialization domain (see Table 4), ordinate: mean error; e_t is expressed in mm, e_r in degree and e_s in mm/pixel.

Table 5

Calibration results: mean error and standard deviation are given for translations, rotations and scaling factors

Calibration method	2D criterion	3D criterion	StradX
Mean error and std. dev. (mm)	0.9594 (0.1137)	0.5454 (0.0952)	1.4874 (0.7293)
Mean error and std. dev. (°)	0.0068 (0.0010)	0.0015 (0.0006)	0.0060 (0.0037)
Mean error and std. dev. (mm/pixel)	0.0014 (0.0004)	0.0021 (0.0003)	0.0080 (0.0091)

The results for translations are in mm, for rotations in degrees, for scaling factors in mm/pixel.

To evaluate the reconstruction accuracy, a 3D ultrasound calibration phantom¹ was scanned. A test object with known volume provided by the manufacturer (6.6162 mm³) was reconstructed. Fig. 14 shows B-scans of the phantom. Therefore, we can compare the volume given by the manufacturer with the estimated volume for different calibration method. Volume estimation is performed using the manual segmentation tool of StradX.

To evaluate the calibration reproducibility, we use a criterion proposed by Prager et al. (1998),

$$\Delta \mathbf{x}^t = \|T_{r1}\mathbf{x}_r - T_{r2}\mathbf{x}_r\|, \quad (9)$$

where \mathbf{x}_r is a corner point of B-scan (this expression includes the scaling factors), T_{r1} and T_{r2} two calibration transformations for the same ultrasound sequence. $\Delta \mathbf{x}^t$ is the reproducibility measure. When $\Delta \mathbf{x}^t$ is close to zero, the reproducibility of the calibration method is good.

5.2.3. Results

Sixteen sequences have been used and some of these have a quite poor quality (see Fig. 8(b)). Table 6 summarizes reproducibility results using the criterion $\Delta \mathbf{x}^t$ and accuracy evaluation with volume measurement

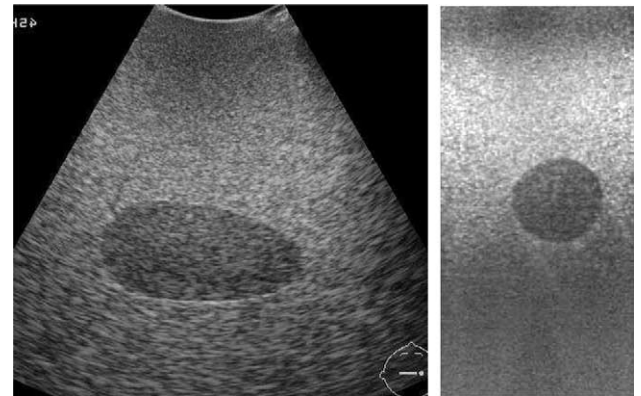


Fig. 14. Representative images of US sequences: phantom B-scans.

method. Concerning volume estimation, the discrepancy between the true volume of the test phantom and its estimated volume is about 1–3% with our method and about 8% with StradX. These results indicate that accurate volume estimation can be obtained with 3D freehand ultrasound system. Moreover, the two proposed formulations are quite equivalent and perform better than the StradX method in terms of reproducibility and precision. 3D approach is more time consuming than 2D approach (around two times longer) but time computation is in all the cases about one to five minutes for tested image sequences on a standard PC (Pentium 3, 866 MHz).

¹ CIRS, model 55. <http://www.cirsinc.com>.

Table 6
Volume measurements and reproducibility measure

Volume	2D criterion	3D criterion	StradX
Mean volume and std. dev. (mm ³)	6.5388 (0.2023)	6.3994 (0.1967)	6.0750 (0.2803)
Reproducibility (Δx^1)	2D criterion	3D criterion	StradX
Mean error and std. dev. (mm)	1.732 (0.427)	2.749 (1.157)	4.109 (0.1534)

The standard deviation is given inside brackets.

6. Conclusion

We present in this paper Confusius, a novel robust, fast and fully automatic method to calibrate a 3D free-hand ultrasound system. Taking into account several criteria to characterize calibration methods, we clearly show a need for robust and automatic calibration procedure. In the proposed method, the calibration parameters are determined by aligning iteratively a plane phantom and a set of points of interest extracted in 2D images. The point extraction step uses the Hough Transform. Robustness is obtained by verifying the consistency of the extraction using 2D and 3D information with the Hough transform and using an outlier rejection process for each B-scan. Robustness and automaticity of the extraction procedure has been compared with StradX software. The proposed method is fully automatic which is a key point to use 3D freehand ultrasound in clinical applications. We investigate a 2D and

yet more time consuming. Our method has been shown to be robust, accurate and fast enough for clinical use.

Acknowledgements

We thank Pierre Darnaud from CHU Rennes, François Chaumette from IRISA-INRIA, Richard Prager and Andrew Gee from Cambridge University for fruitful discussions, and Brittany region for material support. We also thank Wiro Niessen and Marloes Letteboer from the ISI Utrecht for material assistance concerning ultrasound image acquisition.

Appendix A. Rotation matrix

In the Eulerian representation, the matrix R is expressed as follows:

$$\begin{aligned}
 R(\alpha, \beta, \gamma) &= R_z(\alpha) \cdot R_y(\beta) \cdot R_x(\gamma) \\
 &= \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \gamma & -\sin \gamma \\ 0 & \sin \gamma & \cos \gamma \end{pmatrix}, \\
 R(\alpha, \beta, \gamma) &= \begin{pmatrix} \cos \alpha \cos \beta & \cos \alpha \sin \beta \sin \gamma - \sin \alpha \cos \gamma & \cos \alpha \sin \beta \cos \gamma + \sin \alpha \sin \gamma \\ \sin \alpha \cos \beta & \sin \alpha \sin \beta \sin \gamma + \cos \alpha \cos \gamma & \sin \alpha \sin \beta \cos \gamma - \cos \alpha \sin \gamma \\ -\sin \beta & \cos \beta \sin \gamma & \cos \beta \cos \gamma \end{pmatrix}.
 \end{aligned} \tag{A.1}$$

3D formulation of the calibration problem. On the one hand, 2D formulation is well known in visual reconstruction domain where the cost function is defined in images using a image projection model. On the other hand, the 3D formulation using 3D Euclidian distance is more intuitive.

Reproducibility and accuracy of each approach have been evaluated with synthetic and real sequences. The accuracy assessment has been done using a 3D calibration phantom which contains calibrated volumetric test objects. The different formulations have been experimentally compared with StradX software which is a reference in this domain. The performance of the proposed method was significantly better than StradX. 3D approach is slightly more accurate than 2D approach,

Appendix B. Projection of the plane

The plane has to be projected in the image coordinate system. Plane parameters (a, b, c, d) are written as a vector $\mathbf{v} = (a, b, c)$ and a scalar d . First, concerning the projection in the receiver coordinate system, we have

$$\mathbf{v}_r = \mathbf{R}_t^{-1} \mathbf{v}_t,$$

$$d_r = d_t - (\mathbf{R}_t^{-1} \mathbf{v}_t) \cdot \mathbf{t}_t^{-1},$$

where (\mathbf{v}_t, d_t) are the plane parameters in the transmitter coordinate system, and (\mathbf{v}_r, d_r) are the plane parameters in the receiver coordinate system. $T_t = (\mathbf{R}_t, \mathbf{t}_t)$ is the rigid transformation from the receiver to the transmitter coordinate system.

$\mathbf{R}_t^{-1} = \mathbf{R}_t^t$ because \mathbf{R}_t is a orthogonal matrix and $\mathbf{t}_t^{-1} = -\mathbf{R}_t^t \mathbf{t}_t$; thus, this leads to simplifications:

$$\mathbf{v}_r = \mathbf{R}_t^t \mathbf{v}_t,$$

$$d_r = d_t + \mathbf{v}_t \cdot \mathbf{t}_t.$$

In the image coordinate system, we have

$$d_i = d_r + \mathbf{v}_r \cdot \mathbf{t}_r,$$

$$\mathbf{v}_i = \mathbf{S} \mathbf{R}_r^t \mathbf{v}_r,$$

where \mathbf{S} is the scale matrix and (\mathbf{v}_i, d_i) are the plane parameters in the image coordinate system. Finally, we obtain

$$d_i = d_t + \mathbf{v}_t \cdot \mathbf{t}_t + \mathbf{R}_t^t \mathbf{v}_t \cdot \mathbf{t}_r,$$

$$\mathbf{v}_i = \mathbf{S} \mathbf{R}_r^t \mathbf{R}_t^t \mathbf{v}_t.$$

References

- Barry, C.D., Allott, C.P., John, N.W., Mellor, P.M., Arundel, P.A., Thomson, D.S., Waterton, J., 1997. Three-dimensional freehand ultrasound: image reconstruction and volume analysis. *Ultrasound in Medicine and Biology* 23 (8), 1209–1224.
- Blackall, J.M., Rueckert, D., Maurer Jr., C.R., Penney, G.P., Hill, D.L.G., Hawkes, D.J., 2000. An image registration approach to automated calibration for freehand 3D ultrasound. In: *Proceedings of Medical Image Computing and Computer-Assisted Intervention*, pp. 462–471.
- Bouchet, L.G., Meeks, S.L., Goodchild, G., Bova, F.J., Buatti, J.M., Friedman, W.A., 2001. Calibration of three-dimensional ultrasound images for image-guided radiation therapy. *Physics in Medicine and Biology* 46 (2), 559–577.
- Burt, J.P., 1984. *Multiresolution Image Processing and Analysis, Chapter 2: The Pyramid as a Structure for Efficient Computation*. Springer, Berlin, pp. 6–38.
- Carr, J., 1996. *Surface Reconstruction in 3D Medical Imaging*. University of Canterbury, Christchurch, New Zealand.
- Chalmond, B., 2003. *Modeling and Inverse Problem*. Springer, Berlin.
- Deng, J., Sullivan, I.D., Yates, R., Vogel, M., McDonald, D., Linney, A.D., Rodeck, C.H., Anderson, R.H., 2002. Real-time three-dimensional fetal echocardiography – optimal imaging windows. *Ultrasound in Medicine and Biology* 28 (9), 1099–1105.
- Detmer, P.R., Bashein, G., Hodges, T., Beach, K.W., Filer, E.P., Burns, D.H., Strandness Jr., D.E., 1994. 3D ultrasonic image feature localization based on magnetic scanhead tracking: in vitro calibration and validation. *Ultrasound in Medicine and Biology* 20 (9), 923–936.
- Dutt, V., 1995. *Statistical analysis of ultrasound echo envelope*. The Mayo Graduate School.
- Fenster, A., Downey, D.B., Cardinal, H.N., 2001. Three-dimensional ultrasound imaging. *Physics in Medicine and Biology* 46, 67–99.
- Fishler, M.A., Bolles, R.C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* 24 (6), 381–395.
- Gee, A.H., Prager, R.W., Treece, G., Berman, L., 2003a. Engineering a freehand 3D ultrasound system. *Pattern Recognition Letters* 24, 757–777.
- Gee, A.H., Treece, G.M., Prager, R.W., Cash, C.J.C., Berman, L., 2003b. Rapid registration for wide field-of-view freehand 3D ultrasound. Department of Engineering, University of Cambridge, England.
- Holm, S., 1997. Real-time 3D medical ultrasound – signal processing challenges, Norwegian Signal Processing Symposium.
- Hough, P.V.C., 1959. *Machine Analysis of Bubble Chamber Pictures*. International Conference on High Energy Accelerators and Instrumentation, CERN.
- Laveau, S., 1996. *Géométrie d'un système de N caméras. Théorie, estimation et applications*, École polytechnique.
- Lehmann, E.L., 1986. *Testing Statistical Hypotheses*. Wiley, New York.
- Meairs, S., Beyer, J., Hennerici, M., 2000. Reconstruction and visualization of irregularly sampled three- and four-dimensional ultrasound data for cerebrovascular applications. *Ultrasound in Medicine and Biology* 26 (2), 263–272.
- More, J.J., 1978. The Levenberg-Marquardt algorithm: implementation and theory. In: *Proceedings of the Biennial Conference on Numerical Analysis*, pp. 105–116.
- Muratore, D.M., Galloway Jr., R.L., 2001. Beam calibration without a phantom for creating a 3-D freehand ultrasound system. *Ultrasound in Medicine and Biology* 27 (11), 1557–1566.
- Pagoulatos, N., Haynor, D.R., Kim, Y., 2001. A fast calibration method for 3-D tracking of ultrasound images using a spatial localizer. *Ultrasound in Medicine and Biology* 27 (9), 1219–1229.
- Prager, R.W., Rohling, R.N., Gee, A.H., Berman, L., 1998. Rapid calibration for 3-D freehand ultrasound. *Ultrasound in Medicine and Biology* 24 (6), 855–869.
- Prager, R.W., Gee, A.H., Berman, L., 1999. StradX: Real-time acquisition and visualisation of freehand three-dimensional ultrasound. *Medical Image Analysis* 3 (2), 129–140.
- Rousseeuw, P.J., Leroy, A.M., 1987. *Robust Regression and Outlier Detection*. Series in Probability and Mathematical Statistics. Wiley–Interscience, New York.
- Sato, Y., Nakamoto, M., Tamaki, Y., Sasama, T., Sakita, I., Nakajima, Y., Monden, M., Tamura, S., 1998. Image guidance of breast cancer surgery using 3-D ultrasound images and augmented reality visualization. *IEEE Transactions on Medical Imaging* 17 (5), 681–693.
- Triggs, B., McLauchlan, P., Hartley, R., Fitzgibbon, A., 1999. Bundle adjustment: a modern synthesis. In: *Workshop on Vision Algorithms*, pp. 298–372.