



HAL
open science

Towards real-time free-hand biopsy navigation

Clément Beitone, Gaëlle Fiard, Jocelyne Troccaz

► **To cite this version:**

Clément Beitone, Gaëlle Fiard, Jocelyne Troccaz. Towards real-time free-hand biopsy navigation. Medical Physics, 2021, 48 (7), pp.3904-3915. 10.1002/mp.14582 . hal-02997319

HAL Id: hal-02997319

<https://hal.science/hal-02997319>

Submitted on 10 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Towards real-time free-hand biopsy navigation

Clément Beitone* and Jocelyne Troccaz†

*Univ. Grenoble Alpes, CNRS, CHU Grenoble Alpes,
Grenoble INP, TIMC-IMAG, F-38000 Grenoble, France*

Gaëlle Fiard‡

*Univ. Grenoble Alpes, CNRS, CHU Grenoble Alpes,
Grenoble INP, TIMC-IMAG, F-38000 Grenoble, France
Department of Urology, Grenoble Alpes University Hospital, Grenoble, France
(Dated: October 20, 2020)*

Purpose: Performing a transrectal ultrasound (TRUS) prostate biopsy is at the heart of the current prostate cancer detection procedure. With today's 2D live ultrasound (US) imaging equipment, this task remains complex due to the poor visibility of cancerous tissue on TRUS images and the limited anatomical context available in the 2D TRUS plane. This paper presents a rigid 2D/3DUS registration method for navigated prostate biopsy. This allows continuous localization of the biopsy trajectory during the procedure.

Methods: We proposed an organ-based approach to achieve real-time rigid registration without the need for any probe localization device. The registration method combines image similarity and geometric proximity of detected features. Additions to our previous work include a multi-level approach and the use of a rejection rate favouring the best matches. Their aim is to increase the accuracy and time performances. These modifications and their in-depth evaluation on real clinical cases and comparison to this previous work are described. We performed static and dynamic evaluations along biopsy trajectories on a very large amount of data acquired under uncontrolled routine conditions. The computed transforms are compared to a ground truth obtained either from corresponding manually detected fiducials or from an already evaluated registration method.

Results: All results show that the current method outperforms its previous version, both in terms of accuracy (the average error reported here is 12 to 17% smaller depending on the experiment) and processing time (from 20 to 60 times faster compared to the previous implementation). The dynamic registration experiment demonstrates that the method can be successfully used for continuous tracking of the biopsy location w.r.t the prostate at a rate that varies between 5 and 15 Hz.

Conclusions: This work shows that on the fly 2D/3DUS registration can be performed very efficiently on biopsy trajectories. This allows us to plan further improvements in prostate navigation and a clinical transfer.

* clement.beitone@univ-grenoble-alpes.fr; Laboratoire TIMC-IMAG (GMCAO) – Pavillon Taillefer - 38706 La Tronche Cedex - France

† jocelyne.troccaz@univ-grenoble-alpes.fr

‡ gfiard@chu-grenoble.fr

I. INTRODUCTION

Prostate cancer is the second most common cancer in men and the fourth most common cancer overall according to the World Research Cancer Fund (WRCF). Prostate screening tests may include a Prostate Specific Antigen test (PSA) or a Digital Rectal Examination test (DRE), but to determine whether or not prostate cancer is present, biopsy series must be performed. If some samples are positive, they will also be used as raw material to determine the cancer grade (through the Gleason score [1]) and to make decisions about the more appropriate treatment.

As a result, performing transrectal ultrasound (TRUS) biopsy of the prostate is an essential proficiency for urologists. The procedure itself (see Figure IV) is a complex task that requires extensive anatomical and technical knowledge and good hand-eye coordination. In many countries, medical authorities currently recommend the biopsy of twelve regions uniformly distributed over the prostate, along with some specific targeted locations if additional MRI exams have revealed suspicious regions.

With current 2D live ultrasound imaging equipment, the false negative rate for this procedure could reach thirty percent [2]. This is a consequence of the poor visibility of cancerous tissues in TRUS images and the limited anatomical context available in the 2D TRUS plane. This results in a very limited knowledge about the actual position of biopsy cores. Over the years, significant efforts have been made to lower this rate through the accurate localization of biopsy. Accuracy requirements may be related to the notion "clinically significant tumour": 0.5cc is a consensus value accepted by the urology community [3]. This corresponds approximately to a sphere having a radius of 5mm.

As a first attempt for computer-assisted prostate biopsy, some authors proposed to follow the probe during the examination using an external measurement device. We will refer to these approaches as sensor-driven. Various types of devices have been used, from external tracking (optical / magnetic [4]) to inertial units or robotized arms [5] attached to the probe. These approaches generally use a 3D MR reference image acquired during the preoperative stage. This 3D image is rigidly registered to a 3D TRUS image acquired at the start of the intraoperative phase. This first 3D / 3D registration serves as a reference and is continuously updated by "adding" the probe movement during the examination estimated using the tracking device. If these solutions can be performed at a high frequency, none of them takes into account the fact that the prostate can roll over or become deformed due to contact with the head of the probe during the procedure. Some authors [6] report large errors even in the initial MR/TRUS registration phase: up to 10mm. Due to prostate drifts during the operation, these errors will increase and obviously lead to biopsies far from the targets. Finally, most of these methods do not use an appropriate TRUS 3D probe but reconstruct 3D volumes from 2D images using the information given by the tracking device. This reconstruction process is not free from the introduction of prostate movements and deformations resulting from the movement of the probe itself.

To cope with these motions and deformations, techniques based on image fusion have given promising results. We refer to them as organ-driven approaches. One of the first attempts was proposed by Baumann *et al.*[7]. In this work a first stage consists in the elastic registration of preoperative MR images with 3D TRUS images acquired at the start of the intraoperative phase. Unlike the previously introduced methods, this initial MR/TRUS registration is updated using another registration performed during the procedure. It relates the initial 3D TRUS volume to 3D TRUS volumes acquired all along the procedure. This allows to update the target positions based on the real live observed anatomy. For example, in the Urostation (a commercial platform from Koelis SAS), 3D TRUS images are acquired before each biopsy and are registered rigidly or elastically on this first 3D TRUS image. This allows to provide very accurate maps of targeted locations in a reference volume (MRI and TRUS).

However, even with this approach, the physician has no information other than their mental representation between two 3D acquisitions, while they are navigating with the flow of 2D TRUS images. To guide them towards the biopsy locations as efficiently and accurately as possible, an important issue is to link the 2D TRUS flow to the initial 3D image. Refreshing the target position information from 10 to 15 times per second is most likely sufficient to guide the clinician. Some authors have proposed methods for registering the 2D TRUS images to a 3D TRUS reference image. For instance, De Silva *et al.* [8] use normalized cross-correlation as an objective similarity function to rigidly register 2D images onto the 3D reference. A passive arm holding the US probe allows to initialize the unknown transform before registration. The functional is optimized with the Powell Brent method in order to speed up the computation time by avoiding the computation of derivatives. But even with an optimized version of this method and the delegation of parallelized computation to a GPU [9], they had to sub-sample the input images by a factor of 4 to reach the real-time constraint, *i.e.* 10-15 Hz. However, to perform a GPU calculation, the ultrasound platform must be equipped with a specific board which is a strong constraint on current existing hardware. Moreover, this approach also requires a passive arm holding the probe. This potentially makes US probe motion less natural.

To alleviate these constraints, our team [10] has proposed a different organ-based approach to achieve real-time registration on a standard CPU without the need for a passive arm. This method can be seen as a variant of the iterative closest point method (ICP). A hybrid metrics combines image similarity and geometric proximity of detected features. It is used to match points that are finally registered by means of a rigid transform. This first version was tested on a clinical database of more than 100 3D US volumes from 20 patients.

Based on this first evaluation which was already very promising and to which we will compare, further work was necessary to increase accuracy (i.e. decrease average error and variability) while decreasing computation time. We also wanted to evaluate the method on a much larger clinical dataset. In this paper, we present the modifications and optimizations we made to the method and the evaluations we performed on a large dataset composed of more than 900 biopsy volumes from real clinical data. Moreover, our previous work was limited to static 2D/3D registrations. Additional experiments have been conducted to test the ability to track the prostate all along the procedure by testing high frequency 2D/3D registrations along probe trajectories. These dynamic experiments are also reported.

II. MATERIALS AND METHODS

A. Proposed approach

The inputs of the algorithm are a 3D TRUS reference image and a 2D TRUS image. Let us underline that a single 3D US probe allows both acquisition modes. The first 3D reference is acquired just before starting the biopsy series. During the procedure, a new 3D reference is acquired just before the needle insertion for each tissue sampling. The 2D image comes from the US flow acquired when navigating the US probe from one biopsy site to the next. Matching the 2D image on the fly to the latest 3D reference volume will provide information suitable for guiding the physician towards the next biopsy site. Figure 2 illustrates the integration of our method in a routine clinical process, where continuous 2D/3D registrations would be performed between two biopsies. The following sections provide information about the different steps of the registration method. We invite the interested reader to refer to [10] where different solutions and their parametrizations have been studied and are presented in full details.

1. Multi-resolution approach

In order to improve the seminal work proposed by Selmi [10], we have introduced a multi-resolution approach to the original workflow. Multi-resolution approaches are used to obtain more precise and accurate results with a generally better computation time. Regarding registration, this results in finding more distant solutions since a larger search radius can be used at the coarser levels (see section II A 3). In other words, it allows the method to perform a more efficient and extensive search within the solution spaces which size increases with the level. The computational efficiency is also improved as many iterations (inexpensive in computing time) are performed at coarser levels to align the set of points globally and fewer iterations (more expensive in computing time) are needed at finer levels to refine this alignment.

Our algorithm repeats the steps described in the sections II A 2, II A 3 and II A 4 at each resolution level and uses the result (the rigid transformation) of one level as input for the next one. As rather classically done, for a trade-off between efficacy and accuracy, we used a three-level pyramid of images with a halving shrinking factor at every level.

The process performed for each 2D live image is summarized by the flowchart of figure 3.

2. Features detection & Description

As it has been shown by Gillies *et al.* [9], performing 2D/3D registration based on image similarity analysis requires offloading calculations to a GPU unit. To achieve real-time registration on a standard CPU, it is necessary to shift toward a less greedy registration technique. This is why we decided to use, instead of the full voxel matrix, a sparser representation of image: features points.

They correspond to specific locations inside the images which are selected by a feature detector chosen to be reliable and coherent inside a stream of successive images. In our implementation, we use the well-known Harris detector [11]. This simple but robust detector works on both 2D and 3D images. This ability ensures the consistency of the detected points for the next steps of our algorithm. This detector use a gradient-based strategy to detect robust features named "corners".

A descriptor is attached to each of these corners. It contains the coordinates of the corner expressed in the ultrasound probe coordinate frame - we call it feature point. The descriptor also includes grey level information about the point and its neighborhood Ω (the average intensity and a local histogram of the region around the corner). Based on the results of previous experiments [12], this neighborhood is a 5x5 pixels or 5x5x5 voxels region whatever the resolution level where it is computed.

3. Matching

140 The next step consists in matching the sets F_{2D} and F_{3D} of feature points respectively describing the 2D and 3D images. A first transformation is necessary to map the set of feature points describing the current 2D image into the 3D reference space as illustrated in Figure 4. This initial transformation would be given either by the result of the 2D/3D registration of the previous 2D image to the 3D current reference or by the last 3D/3D registration when navigation to the next biopsy site starts. Let us remind that a single probe enables both 2D and 3D acquisitions: 145 without probe motion, the 3D volume is acquired by rotating the 2D sensor from its resting position (most often the volume is symmetric with respect to this resting position its median slice corresponds to the 2D live image position). Calibration information allows representing 2D and 3D acquired information in a single reference frame. Therefore, having registered a volume to the reference (panorama), the median slice of this volume can be easily mapped to the panorama. This gives a good prior for next live image.

150 Pairing combines feature similarity and geometric consistency. Each 2D feature $f_{2D} \in F_{2D}$ has to be paired to a 3D feature $f_{3D} \in F_{3D}$ located in its neighborhood. Let us remind that, during registration, the method has to iteratively determine the neighbourhood of feature points as ICP iteratively computes closest points. To do this efficiently, a KD-Tree is precomputed from the set of 3D features once for all 2D/3D registrations during navigation to the next biopsy site. This first selection based on geometric distance allows to work only on a subset of potential but realistic 155 3D matches for each 2D feature.

Then, the previously selected pairs, based on geometric distance, are matched using image intensity information. For each pair we compute a distance d :

$$d(f_{2D}, f_{3D}) = \alpha_1 d_1 + \alpha_2 d_2 + \alpha_3 d_3, \quad (1)$$

where $\sum_{i=1}^3 \alpha_i = 1$, d_1 compares the mean intensity in the neighborhood Ω of the features, d_2 uses the Bhattacharyya distance which quantifies the similarity of two normalized histograms, d_3 is the sum of squared differences computed in Ω . Finally, for each point of the 2D set, only the best 3D match is kept.

160 In order to improve the registration quality, one solution is to keep among all 2D/3D pairs resulting from the previous stages, only those with the highest similarity values. This selection is based on a fixed rejection rate of the worst pairs and requires implementing an efficient sorting method with respect to the similarity values. Conversely to [10] where this rejection rate was not used because of its computational cost, the current implementation makes efficient use of a sorted multi-set container of C++. Such special container can support multiple elements with the same similarity value while maintaining an insertion rate with a logarithmic complexity. The position of the previously 165 inserted element and the comparison to this element is used to estimate the location of the new element. In this special case the computational complexity can even be constant amortized.

4. Finding the transform

When the feature points are paired, the algorithm proposed by Arun [13] for non iterative least square minimization allows to compute the rigid transform that best matches the pairs. This algorithm aims to minimize the geometric distance between the given pairs. The computed transformation is given in the form of an Euclidean transform T with 6 degrees of freedom. To give more importance to information inside the prostate, we use a variant that allows the assignment of weights. T is computed to minimize

$$\sum_{i=1}^n w_i \times d(a_i, T(b_i))^2 \quad (2)$$

170 where w_i is the weight given to one of the n pairs of matched features $(a_i, b_i) \in F_{3D} \times F_{2D}$. Pairs that have a 3D point inside the pre-segmented prostate are considered more meaningful and are therefore weighted with a higher coefficient than points in the rest of the image. Favouring information inside the prostate combined with the high frequency registration partially overcome the lack of elastic registration.

5. Optimization - Stopping Criterion

175 Subsections II A 3 and II A 4 are repeated until a stable transformation is found. The algorithm stops when the Euclidean distance between the 2D points registered between two successive iterations is lower than a given threshold or after reaching a maximum number of iterations given as input.

B. Evaluation on reference [10] dataset

To assess the presented method, we firstly used the same protocol and database than what was used to evaluate the method proposed by Selmi [10]. This protocol focuses on the evaluation of the method under controlled conditions using data from clinical examinations. Here, the objective is to explore in a systematic way the limits of the presented approach on static 2D/3D registrations. This first evaluation is intended to show, in a direct way, the effect of the additions we made to [10]. The following section presents the details of this benchmark. We are also taking advantage of this first evaluation to carry out a study of the sensitivity of the method to the main parameters (search radius and rejection rate).

1. Patient data set

In this experiment, the database is composed of 160 TRUS volumes for a total of 20 patients. These volumes were acquired during routine clinical biopsy series using two different commercial ultrasound systems (Urostation® and Trinity® from Koelis SAS). The clinical procedures were performed by six different operators from two different hospitals (Grenoble University Hospital and Paris La Pitié University Hospital). Acquired data include for each patient an initial 3D image, named reference or panorama volume ($TRUS_{pano}$) and a set of 3D images ($TRUS_{biopsy}$) acquired all along the biopsy protocol. To simulate 2D images, the median slice of each $TRUS_{biopsy}$ volume is extracted for 2D/3D registration to $TRUS_{pano}$. As mentioned in section II A 3, this median slice corresponds to the live image that would be produced using the 2D mode instead of the 3D one having the US probe in the same position.

2. Ground truth

For all the patients, we used fiducials (visible micro-calcifications) in order to compute the ground truth: the 6D transform that maps the 2D image to the 3D reference. Four to six anatomical landmarks within the prostate (visible both in $TRUS_{biopsy}$ and in $TRUS_{pano}$) were required to create fiducial pairs. Volumes that did not contain enough visible micro-calcifications were discarded from the validation database (and are not included in the 160 mentioned above). These fiducial pairs allow to compute the ground truth T_{GT} (cf. figure 5) using paired-point matching. The fiducial localization error was evaluated to 0.38 mm and the fiducial registration error FRE after pair-point matching was 0.74 mm (± 0.34) (cf. reference [10] for more details).

3. Simulated transforms

To simulate the movements of the TRUS probe or prostate, artificial rotations and translations were generated from the initial T_{GT} to produce $T_{Initial}$ from which 3D/2D registration will be executed. For each volume in the biopsy sequence and each simulated perturbation, $T_{Initial}$ is applied to the biopsy volume (see Figure 6) before extracting the median slice to be registered to the reference volume.

Similarly to [10], two different experiments were conducted. In the first one, rigid transforms are generated by modifying one of the six degrees of freedom at a time. The range of allowed perturbations is systematically explored: rotations (resp. translations) range from -15° to 15° (resp. -15 ; mm to 15 mm). The step is 1° (resp. 1 mm) for rotations (resp. translations). In the second experiment, random rigid transforms are produced by changing the 6 position and orientation parameters at the same time. The explored ranges of perturbations were $[-5^\circ, 5^\circ]$ and $[-7.5$ mm, 7.5 mm]. These ranges are chosen deliberately larger than the values encountered during examinations (as if the method operated in real time on the 2D image stream) in order to show its limits. The obtained transform is noted T_{Reg} .

To evaluate the registration method described in section II A, two TRE values between the reference volume fiducials and the matched biopsy volume fiducials are obtained. The first TRE value is computed before the registration process when $T_{Initial}$ has been applied to the biopsy volume fiducials. The second TRE value is determined after registration when T_{Reg} has been applied to the biopsy volume fiducials. The TRE formula is written has:

$$TRE(mm) = \sqrt{\frac{\sum_{i=1}^{N_f} \|f_{ref}^i - T(f_{biopsy}^i)\|^2}{N_f}} \quad (3)$$

where T stands for the homogeneous transform matrix $T_{Initial}$ or T_{Reg} and $(f_{ref}^i, f_{biopsy}^i)$ corresponds to the N_f pairs of fiducials manually selected in the reference and biopsy volumes. The Figure 6 shows the different transforms involved in the computation of the TRE.

We also recorded for every fiducial f^i and every transform T^j , the residual error on the transformed fiducial

$$Df(mm) = \|f_{ref}^i - T^j(f_{biopsy}^i)\| \quad (4)$$

As explained previously, T_{GT} is computed using a direct least square minimization of the distance between corresponding fiducials from two different volumes. T_{GT} can be interpreted as the best transform that can be computed from the handled experimental data and the residual error can be seen as the lowest bound of registration accuracy achievable for such an experiment. Therefore, a useful information for the evaluation of T_{Reg} is to compare it to T_{GT} . This can be done by measuring the distance between the fiducials of the biopsy image transformed by T_{GT} and by T_{Reg} :

$$\delta_{reg}(mm) = \sqrt{\frac{\sum_{i=1}^{N_f} \|T_{GT}(f_{biopsy}^i) - T_{Reg}(f_{biopsy}^i)\|^2}{N_f}} \quad (5)$$

Thanks to our 3D data, the evaluation is not restrained to bi-dimensional fiducials as it is often the case for evaluation of 2D/3D registration (cf. reference [8]). This might have led to an under estimation of the out-of-plane motion. In this paper, the evaluation is based on true 3D point pairs from two 3D volumes and results give a more accurate view of the performance of the method. Let us remind that the 2D/3D registration result is to be applied to the whole reference volume for navigation purpose towards a target which is most often not included in the 2D image plane; this is why such a 3D evaluation is more adequate.

4. Parameter settings

Except for the search radius and rejection rate for which the settings have been studied specifically (see section III A 1), the other parameters were set according to results of [10]. The maximum number of 2D and 3D feature points is 900 and 40,000 respectively. The size of the neighborhoods is 5x5 pixels in 2D and 5x5x5 voxels in 3D. The weight assigned to feature points is 4 when they are located inside the prostate and 1 otherwise. Regarding the stopping criteria, the convergence threshold is set to 1e-3mm and the maximum number of iterations is 50.

C. Evaluation on a larger dataset

1. Patient data set

In this experiment, again focused on static 2D/3D registration, a larger database was used; it was composed of 900 additional TRUS volumes for a total of 80 additional patients. These volumes were acquired during routine biopsy series with two different commercial US-based guidance systems (Urostation® and Trinity® from Koelis SAS), by urologists from the Grenoble University Hospital. No constraints or special protocols were imposed on physicians to perform these procedures. These data have been declared to the CNIL (French data protection authority) under the reference MR3711140520.

2. Ground truth

Because of the large dataset, it was not feasible to define fiducials on each volume in order to get the ground truth value of the searched transform as previously. Therefore, we have evaluated the registration quality differently. When using the cited guidance systems, in order to map MRI targets to the US data, urologists have first to segment the prostate on the first acquired US volume (named panorama or reference volume). Moreover, during the intraoperative phase, the guidance software provided by the platform, also estimates the 3D/3D rigid transform between the reference volume and volumes acquired whole along the biopsy procedure using a widely evaluated image-based registration [7, 14].

This transform will be considered as the ground truth T_{GT} for this evaluation. In a way rather similar to what was described in section II B 2, the accuracy of the registration method is measured by computing the point-to-point distance between the mesh of the reference volume transformed by T_{GT} and the mesh transformed by another transform: $T_{Initial}$ (before registration) or T_{Reg} (after registration). It is defined as:

$$TRE(mm) = \sqrt{\frac{\sum_{i=1}^N \left\| mesh_{ref}^i - T \left(T_{GT}^{-1} \left(mesh_{ref}^i \right) \right) \right\|^2}{N}} \quad (6)$$

245 where T stands for $T_{Initial}$ and T_{Reg} , $mesh_{ref}^i$ is the i^{th} vertex of the reference mesh, N is the number of vertices of this mesh.

We also applied this mesh-based evaluation to a subset of 100 volumes taken in the dataset of the fiducial-based evaluation (cf II B 1). This enabled the comparison between the fiducial-based evaluation and the mesh-based evaluation.

250

3. Simulated transforms

To simulate only realistic transformations we used the same protocol that has been used for the second experiment of the previous section (see II B 3 and see Figure 5). The random rigid transforms composed to T_{GT} include variations of the 6 position and orientation parameters at the same time to produce $T_{Initial}$. The range used to simulate rotations was $[-5^\circ, 5^\circ]$ and was $[-7.5 \text{ mm}, 7.5 \text{ mm}]$ for the translations.

255

D. Evaluation on trajectory

1. Data generation & Ground truth

As introduced in section I, all previous evaluations concerned 2D/3D static registration. Because the intended use of the method is for navigational assistance during biopsy, additional evaluation along biopsy trajectories of the TRUS probe was mandatory. Capturing live 2D images along real trajectories during clinical procedures and being able to compute a ground truth raises both technical difficulties including the need for probe tracking, clinical additional burden and regulatory issues. This is why we designed a different approach based on a biopsy simulator developed by our team [15] and thoroughly evaluated ([16, 17]). This simulator provides a virtual environment for educational and research purposes. It includes patient data collected from real biopsy series when using the Urostation or Trinity guidance system (Koelis, France). Again, preoperative MRI scans, intraoperative 3D TRUS volumes and meshes are 265 associated with each recorded patient among other information. Using a mock-up of the US probe attached to a haptic device (Phantom Omni, Sensable Devices Inc.), the trainee can train and execute virtual biopsy series. Thanks to the haptic device, the position and orientation of the probe is known and a 2D live US image can be generated from the recorded TRUS volume and displayed to to him/her in real-time.

Using the simulator, real biopsy trajectories can be executed and the associated 2D on-the-fly images can be 270 generated from the $TRUS_{pano}$ volume acquired at the beginning of the examination. Thus, the ground truth T_{GT} is easily known by tracking the probe mock-up.

2. Simulated trajectories

For the purposes of this experiment, we executed two types of stereotyped trajectories. The first one is a classical exploration of the anatomical environment with combined rotations of the probe. This type of trajectory is generally 275 encountered during exams when the urologist moves the probe to find the target area. The second is a pure rotation of the probe. This movement is usually performed when the urologist wants to align the needle to a localized target position, without losing the line of sight he/she has on it [18].

For each trajectory, we tested two different registration strategies. One (noted *sequential*) consists in using the result of the previous 2D/3D registration as input for the current one. In this case, two image processing frequencies are used: 1 out of 2 and 1 out of 5 images. This corresponds approximately to physical acquisition frequencies ranging from 5 to 15 Hz. The other strategy (noted *non-sequential*) is to register all the images independently. In this case, 280 for each new image, the initial transformation is the same as the one that was used for the first image in the sequence.

The method used to evaluate the registration quality is mesh-based as presented in section II C (cf. equation 6).

III. RESULTS

A. Evaluation on reference [10] dataset

The results obtained for the comparative experiments aiming at evaluating our additions to [10] are synthesized in Table I. TRE is given before and after registration for the two methods. For complete comparison to [10], this table also provides the percentage of registrations which resulting TRE is lower than 5 mm.

The results show that we have improved the performance of the method in terms of accuracy (average value and standard deviation) compared to reference [10]. Indeed for the first experiment, Table I shows that the average error is 12% smaller while the standard deviation is divided by 3. Regarding the clinical threshold, more than 80% of the results were already lower than 5 mm. It is now 11% better. Decreasing the variance is also crucial because it shows the repeatability of the method. The results of the second experiment follow the same trend in a more realistic evaluation, since the movement of the hand is never done on a single degree of freedom. Whilst starting farer ("TRE before" is larger), our method results in an improvement of performances : 17% less for the average error, the standard deviation is divided by 1.6 and 11% more registrations result in TRE values below the clinical threshold.

Regarding the current method described in this paper, we performed a statistical analysis of performances. Figure 7 illustrates the distribution of errors (TRE and Df) obtained for the experiments 1 and 2 on the dataset. Table II gives the percentage of errors below the clinical threshold for the conducted experiments. The Shapiro-Wilk test allowed to verify that the Df and TRE distributions were not normal. Therefore, we performed a Wilcoxon statistical test between pre and post registration Df and TRE values. This test shows a very significant difference between the TRE values before and after registration. For experiment 1, conducted on 29,760 samples, the p-value is $2.34e^{-18}$. For experiment 2, on 4,000 samples, the p-value is $3.62e^{-17}$. Regarding Df, Wilcoxon returns a zero value, highlighting a very significant difference between before and after registration error distributions (on respectively 127,000 and 17,000 samples).

For the experiments 1 and 2 the δ_{reg} distance is respectively equal to 3.09 mm (± 0.82) and 2.59 mm (± 0.46). Since the typical voxel size of an ultrasound image is about 0.33 mm on each dimension after reconstruction, this corresponds to a relatively small error (less than 8 voxels). For comparison, let us remind that the FRE corresponding to T_{GT} is 0.74 mm (± 0.34).

Finally, in terms of computing time, the proposed algorithm is able to perform 2D/3D registrations at a frequency that varies between 5 and 15 Hz on a standard CPU (Intel Xeon W2145 with 8 cores @3.74 Ghz). These performances vary according to the value of the parameters and these variations are mostly correlated to the size of the chosen search radii. Indeed, with larger search radii, more points are considered potential matches and need to be evaluated. The effect of parameter settings on the performances are discussed in the next section.

1. Sensitivity to parameter changes

a. Search radius Table III shows the analysis of the registration accuracy when varying the search radius parameter within the KD tree tested by running experiment 2. The impact of this parameter on the method is clearly visible: the larger the radius, the more it will be possible to find a large perturbation. However, this will have a slightly negative impact on the accuracy of the method in case of small perturbations. The underlying reason is that this search introduces more false positives at the beginning of the optimization as more points can be falsely coupled.

b. Rejection rate Table IV shows the impact of the rejection rate (RR) tested by running experiment 2. One can see that this parameter allows to obtain more precise results. Indeed, the results with a non-zero RR are better than zero RR where all the pairs are kept. This is a direct consequence of eliminating the worst pairs that can be considered false positives. Nevertheless, increasing the RR improves the quality of the registration to a certain extent. Beyond this threshold, too many pairs are removed and the optimization fails to find a precise transformation, especially for larger perturbation. This setting also has a positive impact on computation time because fewer pairs are used in the registration process.

From these two sensitivity analyses, we decided to use two different parameter configurations depending on the computing time requirements. The first one was used for the static registration experiments (sections III A and III B): the search radii was (5, 3, 2) mm and the rejection rate was 10%. The second one used for prostate tracking experiments on trajectories (section III C) was chosen to be faster: the search radius was (3, 3, 3) mm and the rejection rate was 10%.

B. Evaluation on a larger dataset

For each of the 900 volumes, 10 perturbations were randomly generated and used 3D/2D registration was performed for these 1000 cases. The corresponding results are summarized in Table V. This experiment demonstrates that scaling up is consistent with the results of the section III A. The average accuracy is almost similar to what was obtained in section III A (3.80mm now versus 3.60) as well as the percentage of registrations below the clinical threshold (86% versus 87). It can also be seen that the standard deviation is lower than in the previous experiment (0.48 versus 1.59). This can be explained by the better quality of the input images which were acquired with a more recent ultrasound platform which has a better image quality.

As for the smaller dataset, we performed a statistical analysis of performances. Figure 8 plots the distribution of TRE errors obtained for this experiment. Again, the Shapiro-Wilk showed that the TRE distributions were not normal and the Wilcoxon statistical test between pre and post registration TRE values gave a very significant difference (p-value = 0) between before and after registration error distributions.

Finally, regarding the comparison of mesh-based and fiducial-based evaluation on the subset of a 100 volumes from the data-set of the fiducial-based evaluation, the average post-registration TRE found was 3.43 mm (± 0.51) when using fiducials and 3.88 mm (± 0.53) when using the meshes.

As our mesh based TRE calculation is based on a single mesh for both images, the meaning of this mesh-based TRE value is rather similar to the distance (δ_{reg}) obtained with the fiducials in the section III A. Indeed, the mesh is transformed with the help of the ground truth transformation, to bring it from the panorama to the biopsy image. The residual error computed on the biopsy fiducials was equal to 2.98 mm (± 0.89) while the mesh evaluation gives 3.88mm. Considering that the vertices of the meshes are much more numerous than fiducials and distributed in a larger volume, these results look consistent.

Therefore, all these results prove that the mesh-based evaluation is consistent to the fiducial-based evaluation and that the results obtained for the 900 volumes can be trusted.

C. Evaluation on trajectory

Although, we tested more trajectories, we selected only one trajectory of each type to illustrate this experiment. All conducted experiments worked very similarly to these two.

The results for the exploration trajectory are given in Figure 9. They demonstrate the advantage of using the *sequential* approach over the *non-sequential* approach. Indeed, even if the *non-sequential* approach is able to register the images (the TRE is always better after registration than before), the efficiency of the approach is limited and shows more limitations when the initial position is far from the solution. The *sequential* approach is more efficient as the previous registration result is used to initialize the current registration. This figure also shows that in the case of fast motions in different directions, corresponding to the peaks of the TRE before registration, the ability to perform registrations at a higher frame rate makes the method more robust. This is shown by the two peaks on the orange curve that are not present on the green curve.

Regarding pure rotation trajectories, the results are presented in Figure 10. This experiment shows very clearly the advantages of the *sequential* registration. When the method is applied in a *non-sequential* way, it is unable to find a solution when the TRE increases very strongly (values higher than 10 mm). In the case of this stereotyped trajectory, the frequency at which the registration is performed is less crucial because the movement is relatively smooth during the trajectory. We can nevertheless notice that the light bump at the beginning of the movement is better smoothed in the case where one image out of two is processed (green curve).

By observing the results from these two scenarios, it seems that the TRE in the sequential approach is even smaller than what was described in the previous sections. However, let us underline that no deformation exists between the 3D volume and the 2D image due to the use of the simulator. Getting the real TRE value will require extensive clinical experiments. This is a further step.

IV. DISCUSSION AND CONCLUSION

In this paper, we have described a method for efficiently registering a live stream of TRUS 2D images to a TRUS 3D reference image. The ultimate objective is to improve the quality and reliability of the prostate biopsy procedure. Our method is based exclusively on the live 2D stream and a 3D image acquired at the beginning of the procedure; no external sensors or robotic arms are required. In addition, our method can operate in near real time on a conventional ultrasound platform and does not require a dedicated GPU unit. These performances are reached because we use

feature points, a subset of key landmarks that are automatically detected on the images, instead of processing all the pixels of these images in a conventional image-to-image registration framework.

385 Although previous work regarding 3D/3D registration proved that the prostate moves and is deformed during the biopsy process, our guess was that these movements and deformations would keep small between two successive 2D US live images processed at a high rate. This is why we implemented a rigid registration method. This allows to keep computational cost compatible with navigation requirements (around 10 to 15 Hz). To make registration more accurate, we also gave more importance to information inside the prostate. Results demonstrate that these choices
390 provide a good trade-off between accuracy and processing time requirements.

We performed three different evaluations on a very large amount of data acquired under uncontrolled routine conditions (more than a hundred patients and more than a thousand volumes in total). We also used two different methodologies to assess our results. They both rely on 3D transformations estimated from 3D data from manual and automatic methods for more reliability. To our knowledge, no other approach in this field has used such a
395 large amount of data from real clinical patients to validate their algorithm. From a technical point of view, most of the other methods proposed in the literature like [8, 9, 18] do not use motorized 3D probes but reconstruct these volumes by combining the data from a localization device and the captured stream of 2D images. This acquisition protocol does not allow them to estimate their ground truth from real 3D data, which in our opinion may introduce an underestimation of out-of-plane displacements.

400 Compared to our previous work [10], we have introduced several changes such as the multi-resolution approach, a rejection ratio of matched feature pairs and a more robust stopping criterion computed using displacement error. Combined with optimizations in the choice of data structures and in the implementation, these improvements have allowed us to achieve better performance both in terms of computation time (initially 0.25 Hz now 5-15Hz) and accuracy (15 % on average with a standard deviation divided by 1.5).

405 In all conducted experiments, TRE after registration is lower than before registration. The difference has been statistically tested when possible (i.e. for fiducial-based evaluation) and is very significant. Coming back to the clinical requirements, let us remind that 5mm is considered as the minimum radius of clinically significant tumors. Therefore, ideally, all biopsy targets should be located with a maximum error of 5 mm. Observing directly the error on fiducials (Df) on a large sample of registration experiments and on many data coming from routine acquisitions
410 allows to be optimistic about the clinical applicability of the method: for experiment 1 (resp. 2), Df after registration is in average equal to 3.23 mm (resp. 3.41 mm) and is below 5mm for 89% (resp. 87%) of the cases. To avoid any mislocation of prostate cores, a final check of the current position with respect to the target can be performed using 3D/3D registration before tissue sampling.

415 The complementary experiments we have conducted allow us to show that registrations can be carried out even more efficiently on trajectories using a sequential approach. Similar trends were found in [8], on phantom data and on a small number of clinical images (16 registrations in total): they report an improvement of the registration accuracy (from 1.63 to 1.81 mm in average). Nevertheless, their method require the use of an arm to initialize the registration algorithm for each acquisition. As a final remark, let us stress the fact that our algorithm, when used in a tracking context (which is our final goal), is not very sensitive to variations in the input parameters. The good performances
420 of our approach relies more on its ability to work at high frequencies as illustrated by our experience with trajectories.

Our next objective will be to integrate this approach on an ultrasonic platform in order to achieve real-time 2D/3D registrations on trajectories performed in clinical routine by an urologist.

ACKNOWLEDGMENTS

425 This work was partly supported by the Investissements d’Avenir programme (Labex CAMI) under reference ANR-11-LABX-0004

The authors have no conflicts to disclose.

-
- [1] R. Hurst, L. Hooper, T. Norat, R. Lau, D. Aune, D. C. Greenwood, R. Vieira, R. Collings, L. J. Harvey, J. A. Sterne, R. Beynon, J. Savović, and S. J. Fairweather-Tait, Selenium and prostate cancer: Systematic review and meta-analysis, 2012.
- 430 [2] T. De Silva, A. Fenster, J. Bax, C. Romagnoli, J. Izawa, J. Samarabandu, and A. D. Ward, Quantification of prostate deformation due to needle insertion during TRUS-guided biopsy: Comparison of hand-held and mechanically stabilized systems, *Medical Physics* (2011).
- [3] F. Cornud, L. Brolis, N. B. Delongchamps, D. Portalez, B. Malavaud, R. Renard-Penna, and P. Mozer, TRUS-MRI image registration: A paradigm shift in the diagnosis of significant prostate cancer, *Abdominal Imaging* (2013).

- 435 [4] S. Xu, J. Kruecker, B. Turkbey, N. Glossop, A. K. Singh, P. Choyke, P. Pinto, and B. J. Wood, Real-time MRI-TRUS
fusion for guidance of targeted prostate biopsies, in *Computer Aided Surgery*, 2008.
- [5] J. Bax, D. Cool, L. Gardi, K. Knight, D. Smith, J. Montreuil, S. Sherebrin, C. Romagnoli, and A. Fenster, Mechanically
assisted 3D ultrasound guided prostate biopsy system, *Medical Physics* (2008).
- 440 [6] V. V. Karnik, A. Fenster, J. Bax, D. W. Cool, L. Gardi, I. Gyacskov, C. Romagnoli, and A. D. Ward, Assessment of image
registration accuracy in three-dimensional transrectal ultrasound guided prostate biopsy, *Medical Physics* (2010).
- [7] M. Baumann, P. Mozer, V. Daanen, and J. Troccaz, Prostate biopsy tracking with deformation estimation, *Medical Image
Analysis* (2012).
- [8] T. De Silva, A. Fenster, D. W. Cool, L. Gardi, C. Romagnoli, J. Samarabandu, and A. D. Ward, 2D-3D rigid registration
to compensate for prostate motion during 3D TRUS-guided biopsy, *Medical Physics* **40** (2013).
- 445 [9] D. J. Gillies, L. Gardi, T. De Silva, S. R. Zhao, and A. Fenster, Real-time registration of 3D to 2D ultrasound images for
image-guided prostate biopsy, *Medical Physics* **44**, 4708–4723 (2017).
- [10] S. Y. Selmi, E. Promayon, and J. Troccaz, Hybrid 2D–3D ultrasound registration for navigated prostate biopsy, *Interna-
tional Journal of Computer Assisted Radiology and Surgery* (2018).
- 450 [11] C. G. Harris and M. Stephens, A combined corner and edge detector., in *Alvey vision conference*, volume 15, pages
10–5244, Citeseer, 1988.
- [12] S. Y. Selmi, E. Promayon, and J. Troccaz, 3D-2D ultrasound feature-based registration for navigated prostate biopsy: A
feasibility study, in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology
Society, EMBS*, 2016.
- 455 [13] K. S. Arun, T. S. Huang, and S. D. Blostein, Least-Squares Fitting of Two 3-D Point Sets, *IEEE Transactions on Pattern
Analysis and Machine Intelligence* (1987).
- [14] M. Oderda et al., Accuracy of elastic fusion biopsy in daily practice: Results of a multicenter study of 2115 patients,
International Journal of Urology (2018).
- [15] S. Y. Selmi, G. Fiard, E. Promayon, L. Vadcard, and J. Troccaz, A virtual reality simulator combining a learning
environment and clinical case database for image-guided prostate biopsy, in *Proceedings of CBMS 2013 - 26th IEEE
International Symposium on Computer-Based Medical Systems*, 2013.
- 460 [16] G. Fiard, S. Y. Selmi, E. Promayon, L. Vadcard, J. L. Descotes, and J. Troccaz, Initial validation of a virtual-reality
learning environment for prostate biopsies: Realism matters!, *Journal of Endourology* (2014).
- [17] G. Fiard, S. Y. Selmi, M. Maignon, A. Bellier, E. Promayon, J. L. Descotes, and J. Troccaz, Validating the Transfer
of Skills Acquired on a Prostate Biopsy Simulator: A Prospective, Randomized, Controlled Study, *Journal of Surgical
Education* (2020).
- 465 [18] D. J. Gillies, L. Gardi, D. Tessier, A. Mercado, and A. Fenster, Ring navigation: an ultrasound-guided technique using real-
time motion compensation for prostate biopsies, in *Medical Imaging 2018: Image-Guided Procedures, Robotic Interventions,
and Modeling*, volume 10576, page 105761P, International Society for Optics and Photonics, 2018.

Captions

- 470 1. TRUS-guided biopsy procedure
2. Workflow of the biopsy procedure
3. Main steps of a 2D/3D registration of the live 2D image to the 3D reference image. 3D features are extracted once before series of 2D/3D registrations
- 475 4. Features iterative matching. In blue (resp. orange: the 3D (resp. 2D) feature points set from the reference US volume (resp. 2D on-the-fly US image) for a given stage of the iterative registration
5. Computing the ground truth: T_{GT} is computed from visible fiducials in two US volumes and allows to evaluate the accuracy of registering the median slice extracted in one volume to the other volume.
6. Transformations used by the method. T_{GT} is the ground truth transformation from the biopsy volume and the panorama, $T_{Initial}$ is the initial transformation applied on the biopsy volume and T_{Reg} is the output of the registration algorithm.
- 480 7. Distribution of Df (top) and TRE (bottom) errors for experiments 1 (left) and 2 (right).
8. Distribution of TRE errors for the experiments on the large dataset.
9. Results of the registration for the trajectory simulating an exploration around a given position.
10. Results of the registration for the trajectory simulating a pure rotation around the probe axis.

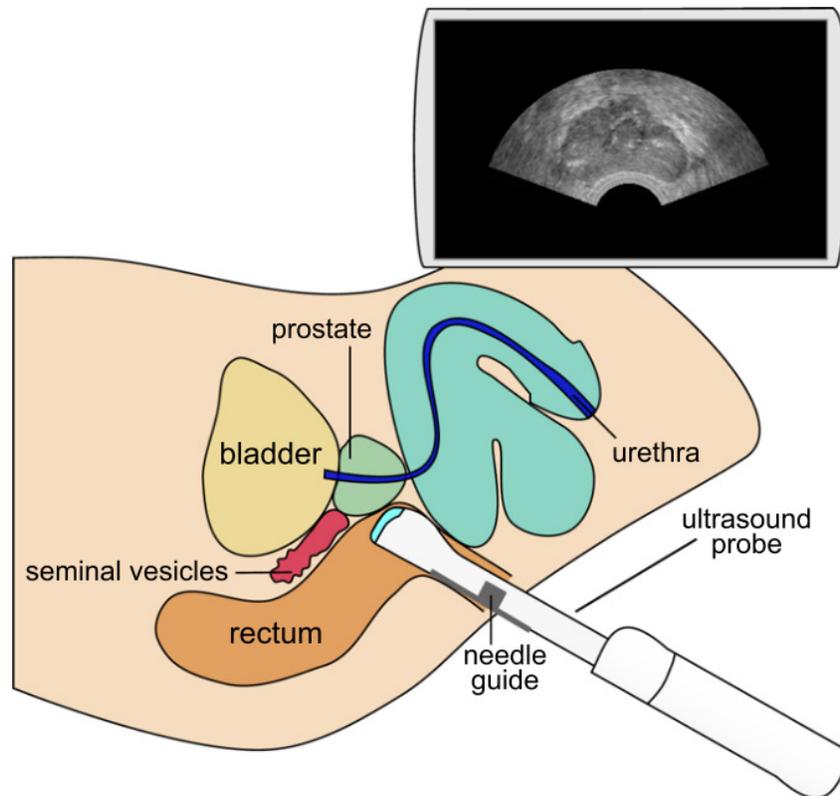


Figure 1. TRUS-guided biopsy procedure

Experiment 1	TRE before	TRE after	TRE after < 5 mm	Df before	Df after	Df after < 5 mm
Reference [10]	8.06 mm (± 4.83)	3.91 mm (± 3.22)	81%	*	*	*
Proposed	8.06 mm (± 4.83)	3.44 mm (± 1.11)	92%	7.98 mm (± 5.12)	3.23 mm (± 1.88)	89%
Experiment 2						
Reference [10]	7.48 mm (± 2.25)	4.37 mm (± 2.62)	76%	*	*	*
Proposed	8.02 mm (± 1.95)	3.60 mm (± 1.59)	87%	7.94 mm (± 3.21)	3.41 mm (± 1.91)	87%

Table I. Summary of the results obtained for the two experiments (section II B) with registration initializations and datasets similar to [10] - reminder of [10] performances. Df was not computed in [10] and is replaced by *

Experiment 1		
	Before	After
Df	44%	89%
TRE	40%	92%
Experiment 2		
Df	19%	87%
TRE	20%	87%

Table II. Percentage of errors below the clinical threshold (5mm) before and after registration.

Search radii per level	TRE (std)
(4., 2., 1.)	4.17 mm (± 1.97)
(3., 3., 3.)	3.54 mm (± 1.32)
(5., 3., 2.)	3.48 mm (± 1.04)

Table III. Impact of the search radius in mm (rejection rate: 0., maximum iteration number: 50)

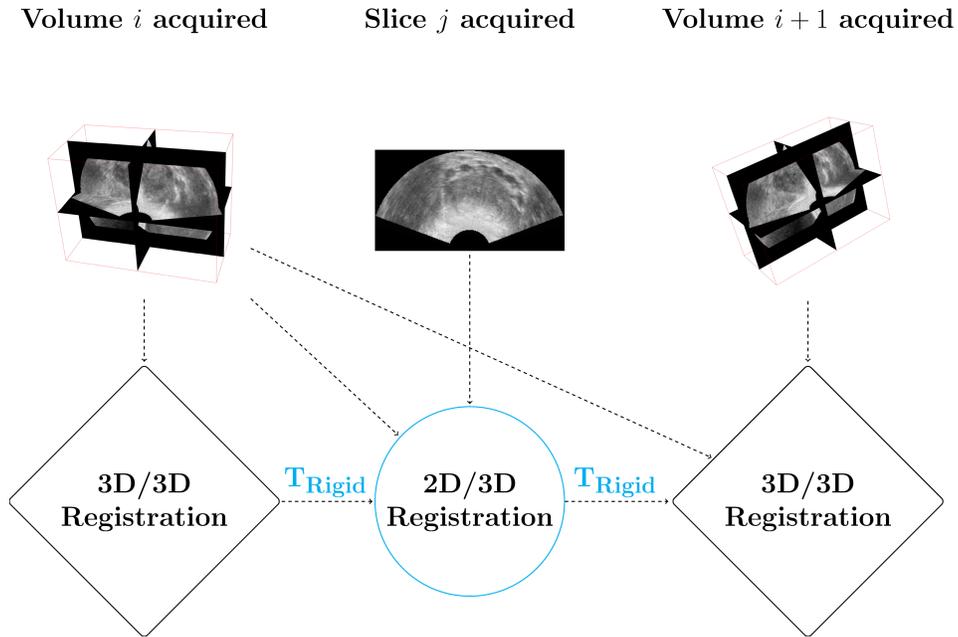


Figure 2. Workflow of the biopsy procedure

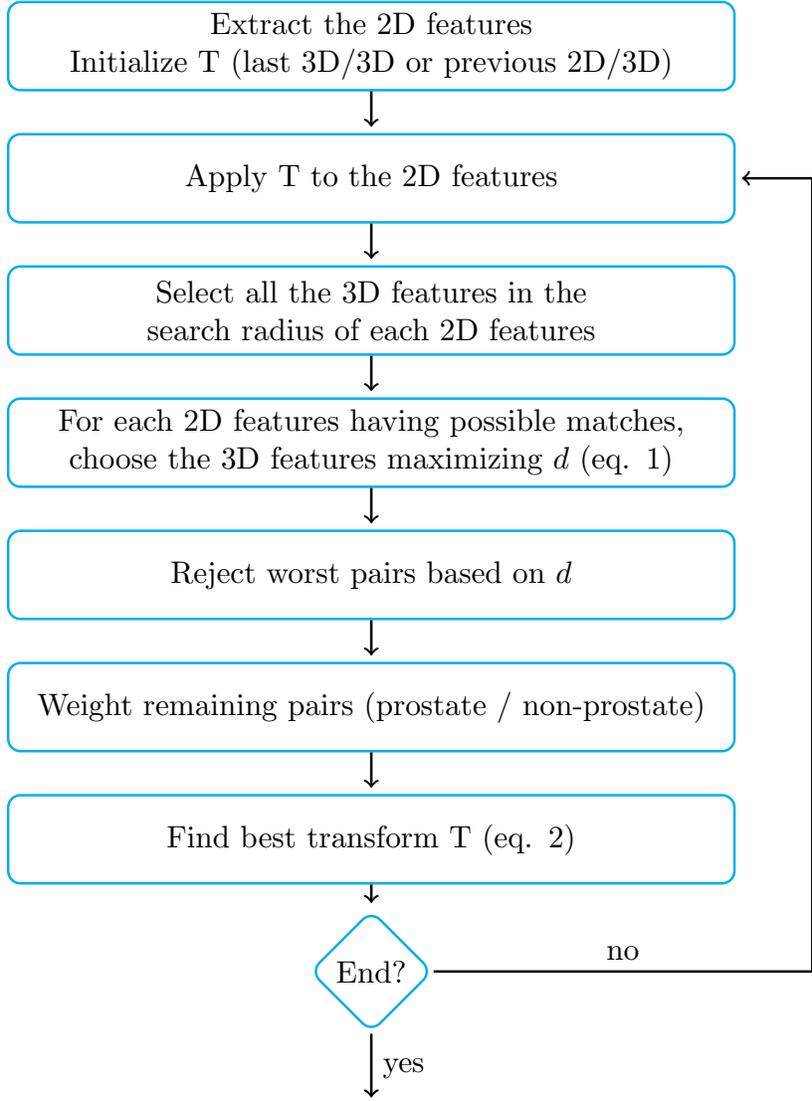


Figure 3. Main steps of a 2D/3D registration of the live 2D image to the 3D reference image. 3D features are extracted once before series of 2D/3D registrations

Rejection rate per level	TRE (std)
(0.3, 0.3, 0.3)	3.70 mm (± 1.36)
(0.3, 0.2, 0.1)	3.60 mm (± 1.30)
(0., 0., 0.)	3.48 mm (± 1.04)
(0.1, 0.1, 0.1)	3.44 mm (± 1.11)

Table IV. Impact of the rejection rate (maximum iteration number: 50, search radii in mm: (5., 3., 2.))

Experiment	TRE before (mm)	TRE after (mm)	TRE after < 5 mm
Proposed	8.56 (± 4.83)	3.80 (± 0.48)	89%

Table V. TRE errors for experiments on the large dataset.

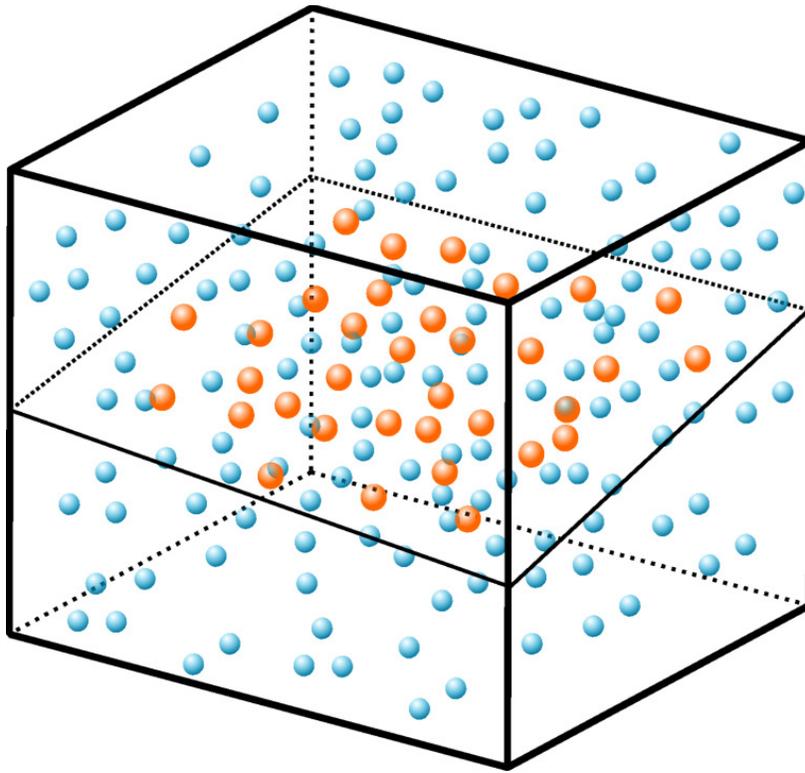


Figure 4. Features iterative matching. In blue (resp. orange: the 3D (resp. 2D) feature points set from the reference US volume (resp. 2D on-the-fly US image) for a given stage of the iterative registration

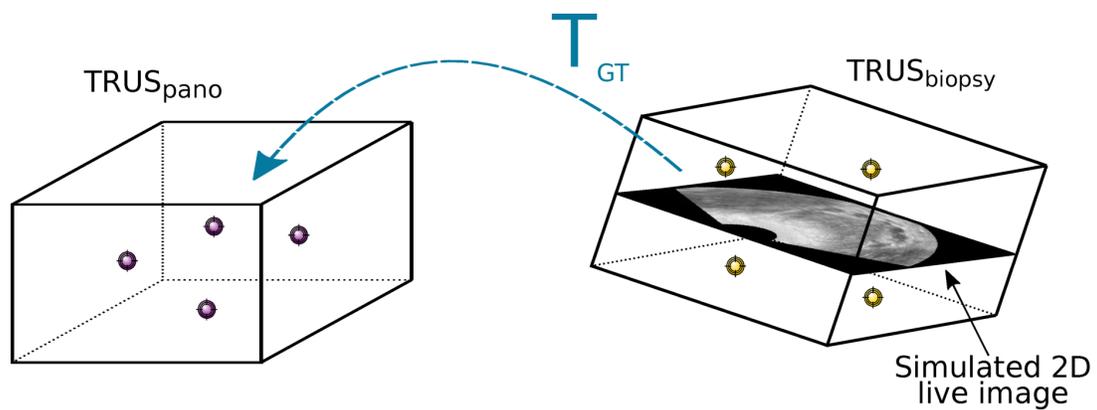


Figure 5. Computing the ground truth: T_{GT} is computed from visible fiducials in two US volumes and allows to evaluate the accuracy of registering the median slice extracted in one volume to the other volume.

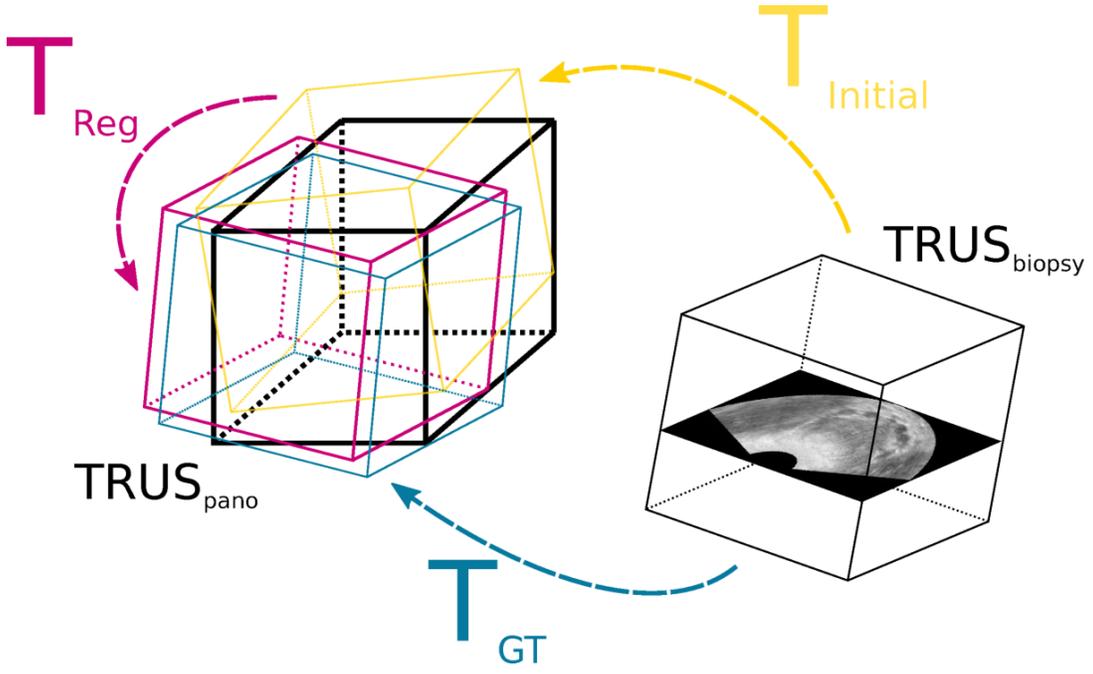


Figure 6. Transformations used by the method. T_{GT} is the ground truth transformation from the biopsy volume and the panorama, $T_{Initial}$ is the initial transformation applied on the biopsy volume and T_{Reg} is the output of the registration algorithm.

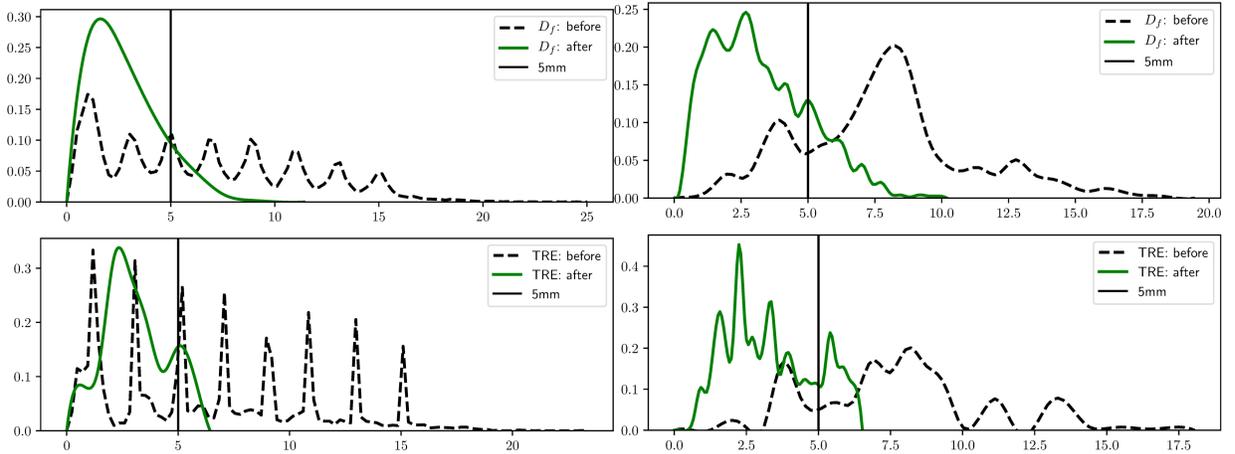


Figure 7. Distribution of D_f (top) and TRE (bottom) errors for experiments 1 (left) and 2 (right).

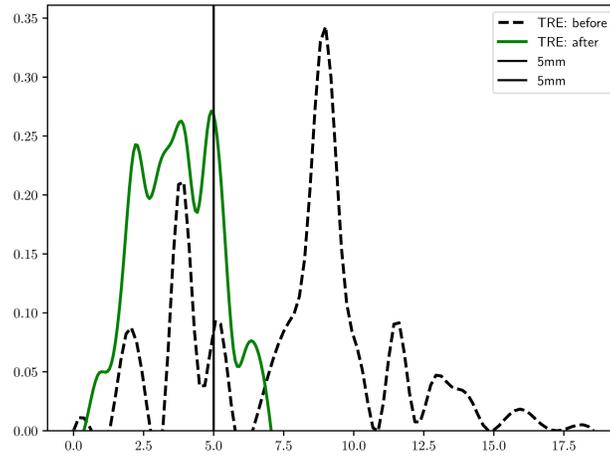


Figure 8. Distribution of TRE errors for the experiments on the large dataset.

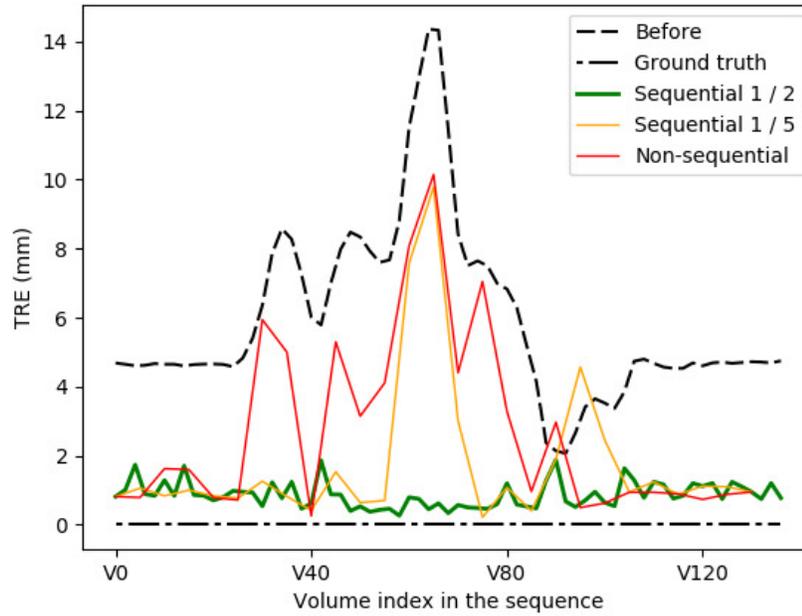


Figure 9. Results of the registration for the trajectory simulating an exploration around a given position.

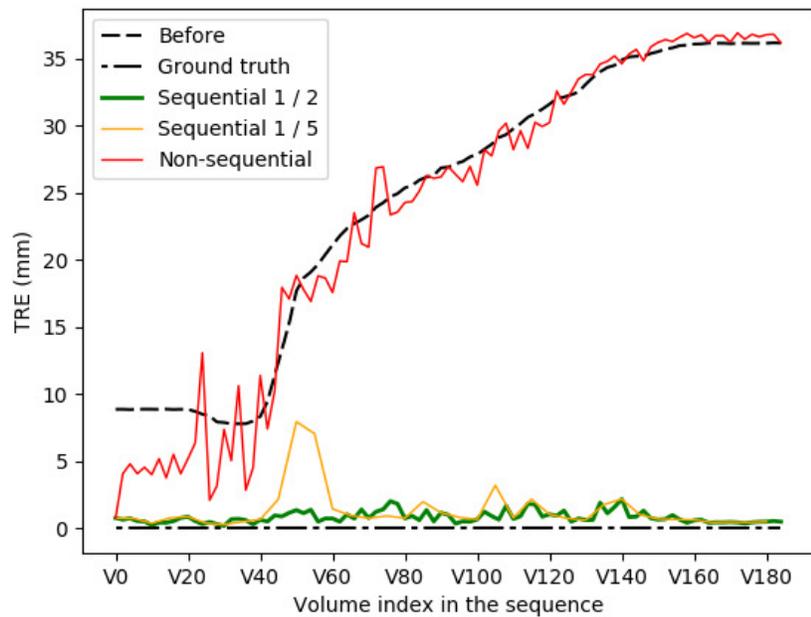


Figure 10. Results of the registration for the trajectory simulating a pure rotation around the probe axis.