

Intensity-based registration of freehand 3D ultrasound and CT-scan images of the kidney

Antoine Leroy · Pierre Mozer · Yohan Payan ·
Jocelyne Troccaz

Received: 29 November 2006 / Accepted: 6 March 2007
© CARS 2007

Abstract

Objectives This paper presents a method to register a pre-operative computed-tomography (CT) volume to a sparse set of intra-operative ultra-sound (US) slices. In the context of percutaneous renal puncture, the aim is to transfer planning information to an intra-operative coordinate system.

Materials and methods The spatial position of the US slices is measured by optically localizing a calibrated probe. Assuming the reproducibility of kidney motion during breathing, and no deformation of the organ, the method consists in optimizing a rigid 6 degree of freedom transform by evaluating at each step the similarity between the set of US images and the CT volume. The correlation between CT and US images being naturally rather poor, the images were pre-processed in order to increase their similarity. Among the similarity measures formerly studied in the context of medical image registration, correlation ratio turned out to be one of the most accurate and appropriate, particularly with the chosen non-derivative minimization scheme, namely Powell-Brent's. The resulting matching transforms are compared

to a standard rigid surface registration involving segmentation, regarding both accuracy and repeatability.

Results The obtained results are presented and discussed.

Keywords Computer-assisted surgery · Kidney · Intensity-based registration · CT-scanner · Freehand ultrasound

Introduction

Clinical and scientific context

Percutaneous renal puncture (PRP) is a common surgical procedure that precedes the extraction of renal stones through a nephroscope. It consists in percutaneously inserting a long and thin needle into one of the narrow calices of the kidney under intra-operative fluoroscopic or ultrasound control. This procedure is an alternative to lithotripsy or ureteroscopy when the stones are too numerous, too big or hardly reachable.

The pre-operative imaging modality is computed-tomography (CT). An early acquisition of the kidneys after injection of a contrast product allows seeing clearly the kidney parenchyma, whereas a late acquisition enhances the tree-like internal cavities (see Fig. 1). Intra-operatively, either fluoroscopy or ultrasound imaging is used for target visualization. So the surgeon has to perform a mental matching between the static horizontal slices from CT and the dynamic intra-operative images to find the targeted cavity, and to decide which way is best to reach it percutaneously. The interviewed surgeons believe that PRP accuracy could benefit from computer assistance, to cope with the lack of visibility, the poor quality of ultrasound images, the irradiating nature of X-Rays and with the difficulty to mentally represent the 3D localization of the small target inside the kidney from those imaging modalities.

This work was partly supported by PRAXIM-Medivision, by the French National Agency for Research and Technology (ANRT) and by the non-profit organization "Association Française d'Urologie".

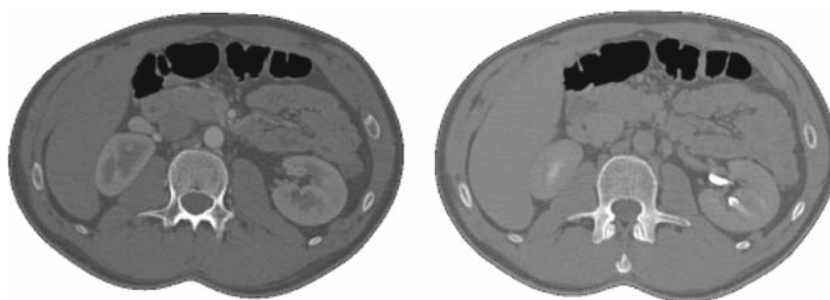
A. Leroy · P. Mozer · Y. Payan · J. Troccaz (✉)
TIMC Lab – IN3S, Faculté de Médecine, Domaine de la Merci,
38706 La Tronche cedex, France
e-mail: Jocelyne.Troccaz@imag.fr

A. Leroy
e-mail: antoine.leroy@koelis.com

P. Mozer
e-mail: pierre.mozer@psl.ap-hop-paris.fr

Y. Payan
e-mail: Yohan.Payan@imag.fr

Fig. 1 CT early (*left*) and late (*right*) acquisitions



Therefore, a feasibility study on computer-assisted PRP was previously carried out [1,2], in which the kidney surface, segmented from CT and localized ultrasound (US) images, was registered using a surface registration algorithm, enabling navigation of the needle. The study showed satisfying results briefly reported in section “Overview of the Method”, however, it required the intra-operative segmentation of US images, in which automation is very difficult and for which an interaction with the surgeon is not practicable for a clinical use.

Therefore, automated CT/US registration was investigated. It was decided for the present study to propose and evaluate a voxel-based registration algorithm, in order to avoid segmentation steps and to minimize user intervention. The developed method has been compared to the surface-based registration of the same data.

State of the art

Voxel-based registration methods have been thoroughly studied since 1995. Every method proposes a similarity measure and a cost minimization algorithm. Wells et al. [3] first introduced Mutual Information (MI) combined with histogram windowing and a gradient descent algorithm. Maes et al. [4] presented an interesting combination of MI and Powell-Brent (PB) search strategy. Maes et al. [5] compared various searches and multi-resolution strategies, and showed that PB was efficient with image subsampling. Jenkinson and Smith [6], Studholme et al. [7] and Sarrut and Miguet [8] made thorough comparisons of different functional and statistical similarity measures. Pluim et al. [9] and more recently [10] proposed to add gradient information for similarity computation.

Although these studies constitute the base of our research, none of them specifically involves US images. We therefore focused on the works of Roche et al. [11,12], who registered 3D US of the brain with MRI, and Penney et al. [13], who registered 2.5D freehand US of the liver with MRI. The theoretical difficulty in registering CT and US is that the former gives information on tissues intensity, whereas the latter contains a speckled image of their boundaries. So a complex similarity measure and specific image preprocessing must be chosen.

In that context, Penney and colleagues’ work [13] is based on probability maps: those maps associated with the data to be registered correspond to the probability of each pixel/voxel to belong to the hepatic tree or to the parenchyma; those transformed data built from a learning phase being mono-modal can thus be registered using a simple normalized cross-correlation. This approach is a very elegant trade-off between contour-based registration requiring the accurate segmentation of data and intensity-based ones that measure image mutual dependency. It benefits from the fact that the hepatic tree is quite visible and different from the parenchyma. Such an approach is not easily applicable to kidney images. Our choice was therefore to process CT and US images before registration to make them more similar whilst keeping in the multi-modality framework. However, the method we propose shows several similarities to that presented by Penney and colleagues: rigidity of the organ assumed, breathing cycle taken into account, specific pre-processing of images including acoustic shadow removal, non-derivative iterative optimization scheme, manual point-to-point pre-registration.

Overview of the method

This paper introduces a method to automatically rigidly register localized US images of the kidney onto a high-quality abdominal CT volume. This method is intended to be integrated in a navigation system (see [1,2]) where pre-operative planning is performed on two registered CT scans; US data are acquired intra-operatively during patient apnea, and then registered to CT data; finally navigation of the puncture needle is performed in the same breathing state than US data acquisition. The US probe and needle are localized using a passive optical device (Polaris[®] from NDI).

The selected intensity-based registration method uses image preprocessing in both modalities, Powell-Brent method as a search strategy, and correlation ratio (CR) as a similarity measure. Tri-linear interpolation is used when superimposing one data set on to the other before CR computation.

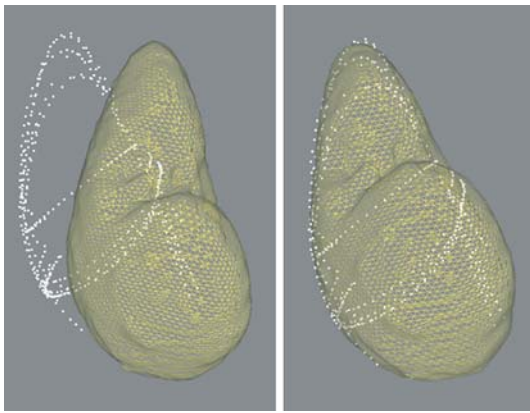


Fig. 2 Rigid surface registration of US and CT segmented kidney (*left* before registration; *right* after registration)

Preliminary tests have been carried out and are compared to a “Bronze Standard”¹ transform (BS)—as defined in [14] based on 3D/3D surface matching [1,2] (Fig. 2) using octree distance maps and Levenberg–Marquardt minimization.

The assessment of surface-based registration on real data was performed in different ways: by testing reproducibility from different initial attitudes (initial errors up to ± 30 mm, $\pm 30^\circ$); by testing coherency in a registration circuit (early-injection CT to US, late-injection CT to US, early-injection CT to late-injection CT); and by computing residual distance between registered data sets on surface, centroid and axis of the kidney. For initial attitudes having their six parameters smaller or equal to 20 (mm or $^\circ$), the convergence and coherency are very good and the average residual distances are respectively 1 mm (surface), 1.4 mm (centroid) and 3.6° (axis). Including initial errors up to 30 (mm and $^\circ$) gives in an average 1.6 mm, 3.8 mm and 8.4° .

Additional tests on a phantom including needle guidance and comparison from post-operative CT data between planned trajectories and inserted needles gave an average error of 4.7 mm measured at the needle tip; this error cumulates the system error and the additional errors due to the segmentation and registration processing steps of post-operative CT data. The analysis of these experiments showed that a major part of the global error is due to needle flexion during insertion in the phantom (let us remind that the use of the optical localizer makes the direct tracking of the needle tip inside the body impossible).

This surface registration serves as a standard for comparison to the method presented in the paper.

¹ By “Bronze standard” we mean a method thoroughly evaluated and that can be considered as a reference but which is not supposed to be the ground truth as is a “gold standard” (based for instance on the use of artificial or anatomical landmarks).

Reproducibility of kidney movements

The approach assumes that the kidney is in the same position when getting intra-operative US data and for percutaneous puncture; both being performed during apneas of the patient. As the kidney moves with respiration, the problem is therefore to verify how much this motion is reproducible from one breathing cycle to another.

Langen and Jones [15] report several clinical studies of kidney motion using different modalities: CT, US, MRI, X-rays. All these studies agree on the fact that the motions amplitude of the kidney is in average between 1 and 2 cm during normal breathing but no information is given about the repeatability of these motions. Schwarz and colleagues [16] report a study on 14 patients from fast MRI data, the measurements being made on two acquisition planes: in average the kidney motion is about 17 mm and its reproducibility is 3 mm. However, we believe that measurements on two low-resolution MRI slices cannot give accurate values on the 3D motion of the kidney.

To complement this knowledge and verify the assumption concerning the reproducibility of kidney movements, experimental work was carried out using localized US images. Static and dynamic acquisitions were made on 11 volunteers. The static acquisitions consisted in sweeping the localized US probe over the kidney whilst each volunteer was carrying out three experiments of full inhales (50 US images recorded per experiment) and two experiments of full exhales (30 US images per experiment). For each experiment, US images were manually segmented leading to a 3D set of data points located at the surface of the kidney. By registering two sets of surface points coming from two different experiments using the rigid surface registration algorithm described in the previous section, it was possible to estimate the kidney displacement between inhale and exhale (one inhale experiment compared to one exhale experiment) as well as the reproducibility of the kidney position for each breathing cycle (one inhale—resp. exhale—experiment compared to another one inhale—resp. exhale—experiment).

To complete the estimation of the kidney displacement between inhale and exhale provided by these static acquisitions, dynamic acquisitions were carried out consisting in aligning the ultrasound plane to the main plane of kidney motion so as to see the longitudinal axis of the kidney, and in acquiring its motion during five breathing cycles (leading to the record and segmentation of 100 dynamic US images per subject).

To make the kidney motion easier to visualize and to interpret, the displacements of two points—the centroid G and the pole P —as well as the inclination α of the principal axis of the organ were measured from those acquisitions (Fig. 3).

Table 1 summarizes the results obtained for the estimations of the kidney displacement during one breathing cycle.

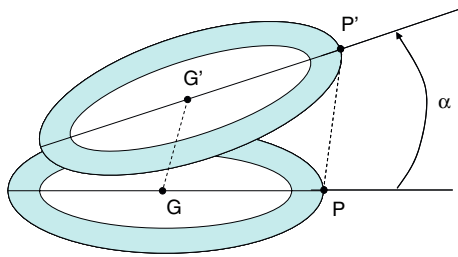


Fig. 3 Kidney motion parameters. GG' and PP' are the displacements of the centroid (G) and the pole (P) of the organ; α is the inclination from exhale to inhale position

Table 1 Measured amplitude of the kidney motion for 11 subjects (in mm and degrees)

	Mean	Std.dev	Min	Max
PP'	30.8	14.1	7.8	55.5
GG'	30.1	14.7	10.1	60.0
α	11.6	3.8	6.5	17.5

Table 2 Reproducibility of centroid position for each breathing cycle (in mm)

	Mean	Std.dev	Min	Max
ΔG	5.9	3.9	1.3	15.9

Concerning the evaluation of the reproducibility of the kidney position for each inhale or exhale (Table 2), it was decided to focus on GG' distance as α is rather small and PP' values are rather similar to GG' ones.

In terms of position repeatability the worst results were obtained for two subjects having the largest index of body mass and for which the ultrasound acquisition and processing turned out to be quite difficult. The best results were obtained for a sportsman having very good control of his apneas, with a measured average reproducibility of 1.8 mm. These results are coherent with published studies and encourage our hypothesis that the kidney motion is reproducible. In operative conditions, using a breathing machine, the results should be improved. A study on patients is envisioned, as well as the use of localized 3D ultrasound.

Material

CT images

CT-scan images are organized in a stack of parallel horizontal slices. Every voxel of the CT volume conveys information on the X-ray intensity at a given position in space. So, homogeneous regions or organs in the body appear as homogeneous grey-level regions in the slices. The volumes we worked on have a typical size of $512 \times 512 \times 300$ voxels, with an interslice space of 0.625 mm and a slice thickness of 1.25 mm.

Thanks to dedicated software (AnalyzeTM), the data were first re-sampled so that the voxel be cubic, with a 0.625 mm width.

As earlier mentioned, more specifically to the application, two different CT scans are available: early- and late-acquisition scans. In the early scan, the contrast product fills the highly vascularized kidney parenchyma; the kidney external shape is thus more easily segmentable, thanks to a homogeneous higher Hounsfield value. In the late scan, after a few minutes, the contrast product flows into the internal kidney cavities, thereby providing an easy delineation in the images of a bright tree-like structure inside the organ (see Fig. 4).

The kidneys are first semi-automatically segmented using Nabl's 3D watershed (from Generic Vision Inc.) in the two CT scans; and then internal cavities may be extracted for planning from the late scan by thresholding. The registration of the kidney from early to late acquisition allows the surgeon to perform a 3D diagnosis and pre-operative planning, with both the information of the kidney's external shape (used for computation of the reference registration) and the cavities (where the target is defined). This mono-modal registration is performed using a classical chamfer rigid matching based on the kidney surface visible on both CT acquisitions.

In the current study, the late-acquisition CT scan is selected for registration to intra-operative data as it showed a better correlation with the US images than the early-acquisition CT.

US images

Ultrasound images convey information on the acoustic interfaces between the organs. They are sometimes called "gradient images" for that reason. The acoustic waves that propagate through the soft tissues emit a reflected echo at each interface and lose intensity with increasing distance to the probe. On the contrary, structures with a higher reflection coefficient (bone, air, renal stone, etc.) tend to reflect

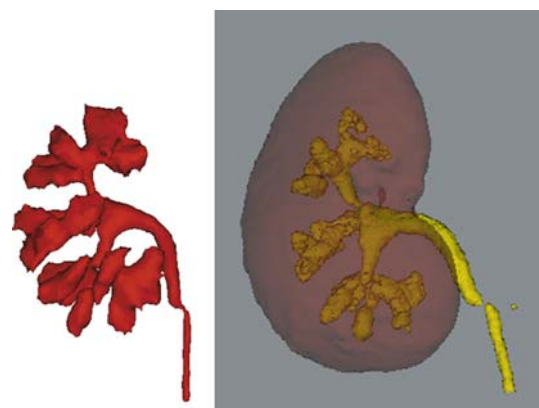


Fig. 4 Right kidney with its internal cavities. (Left) internal cavities segmented from late CT acquisition and (right) repositioned with respect to the kidney external surface segmented from the early CT acquisition

the totality of the signal, thus creating an “acoustic shadow” behind them wrt the probe. US images are known to be highly noisy, the noise called “speckle” being generated by the numerous micro-reflections over the swept area.

Not only the interfaces between regions are represented in US images, but also the regions themselves, to a certain extent, because every region has its own internal cell organization that generates a specific speckle signal. This property should make the comparison between CT and US images feasible.

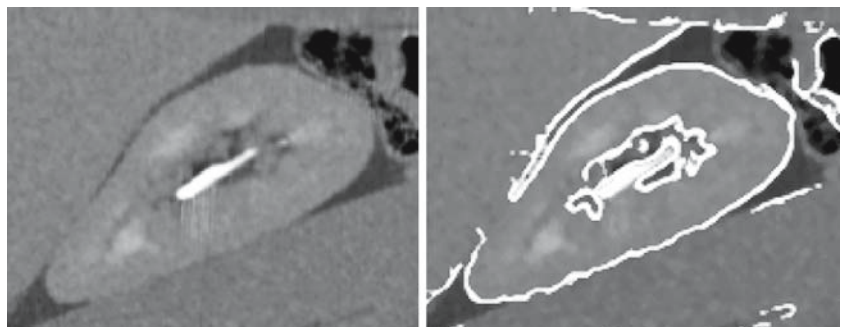
The US images are calibrated and localized in space with the help of an optical localizer. Calibration is performed using Lango’s method [17]. The acquisition of dozens of localized images allows building a sparse set of US voxels in space, that we aim at registering with the dense volume of CT data. The images acquired through an analog video board have a typical size of 768×576 pixels, the pixel being $0.3 \times 0.3 \text{ mm}^2$ (US device: Hitachi-EUB-405, 3.5 and 7.5 MHz probes).

Method

Intensity-based registration is an iterative method to search for an optimal transform T between two sets of data V_1 and V_2 —in this case V_{CT} and V_{US} ; the optimality being defined by a similarity measure between the pixels/voxels of the two data sets.

As previously introduced, the method starts with a pre-processing of the images to make them more similar. Due to the presence of local minima in the optimization process, pre-registration is necessary for robust initialization of T ; the initial value of T , called initial attitude (IA) is computed by a landmark registration. Because the environment of the kidney may be deformed (liver especially), a rigid registration on the whole image would be inappropriate. Therefore, the similarity is computed in a special region of interest (ROI), delineated by the user “around” the kidney. The ROI is defined as the smallest parallelepiped parallel to the US image axes that includes the kidney and its boundaries on the image. As regards the registration process itself, a similarity measure among those available, and an adequate optimization scheme must be chosen.

Fig. 5 Oblique CT kidney after preprocessing: low-pass and contour delineation. (*Left*) before; (*right*) after



This may be summarized by the following equation:

$$T_{\text{opt}} = \arg \max_T (\text{similarity}_{\text{ROI}}(V_{CT}, T(V_{US})))$$

where V_{CT} is the CT data volume and $V_{US} = \{S_1, S_2, \dots, S_k\}$ the set of localized US slices. Let us remark that US data are mapped to CT ones since the CT volume is dense whilst US slices are sparse data.

Preprocessing

CT-scan images

The goal of this stage is to highlight the major boundaries in CT in order to increase the similarity with the “gradient” US images.

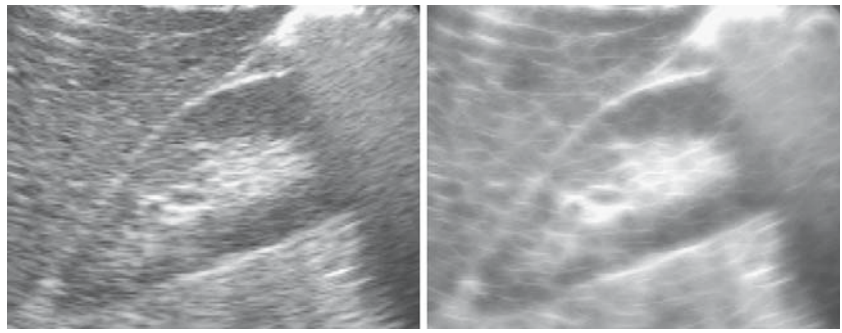
Penney and colleagues [13] basically transformed the blurred MRI and US images of the liver into maps giving the probability of a pixel to be liver tissue or vessel lumen. However, as this process requires a manual thresholding of the MRI, and as the segmentation of the kidney parenchyma is not reduced to a binarization process, especially in the US images, this technique cannot be applied to the kidney problem. Roche et al. [12] proposed the combined use of the MRI image and its derivative. Again, it was decided not to use this approach because of the complexity of the bivariate correlation method and because of the observed chaotic correlation between CT gradient and US.

The selected approach is the following: the preprocessing of a CT slice consists in the superimposition of a median blur and a bi-directional Sobel gradient filter. After contrasting and thresholding the gradient image, the largest connected components are kept and superimposed on to the blurred image (Fig. 5). Sobel filter is adequate here as it creates fairly wide boundaries, whose contribution in the similarity measure will lead to deeper minima in the cost function.

US images

Speckle removal. US images are known to be blurred by speckle noise. Still, the kidney, due to its small size and echogenic capsule, can be well and fully visualized through anterior access. The aim of preprocessing US images is to

Fig. 6 Preprocessing of a US slice using sticks. Most of the speckle is removed while boundaries are preserved. (Left) before; (right) after



reduce the speckle, while preserving the echogenic boundaries of the organ.

The Sticks filter, which used for speckle removal was theoretically justified by Czerwinski et al. [18], was experimentally compared as regards registration to two other filters taken from the literature (median and “salt and pepper” Crimmins’). We finally chose the Sticks filter because, while performing a fairly efficient speckle removal, it kept the bright and “high frequency” boundary information of the kidney and this resulted in improved registration convergence. This filter consists, over an “ $n \times n$ ” convolution window, in computing the maximum of the average grey level over each oblique stick in the window. An example result is shown in Fig. 6.

Shadow removal. Because they convey no relevant information for intensity-based registration, acoustic shadows are removed in the US images, as proposed in [13]. The shadows are produced by large interfaces like the ribs or the colon that quasi-totally reflect the acoustic signal, the remaining US waves decreasing in the distance in an exponential way. In other words, the shadow generally follows a maximum of the intensity profile along a US scan-line. Shadow removal is therefore based on the correlation between a US scan-line profile and a heuristic exponential function. Figure 7 shows

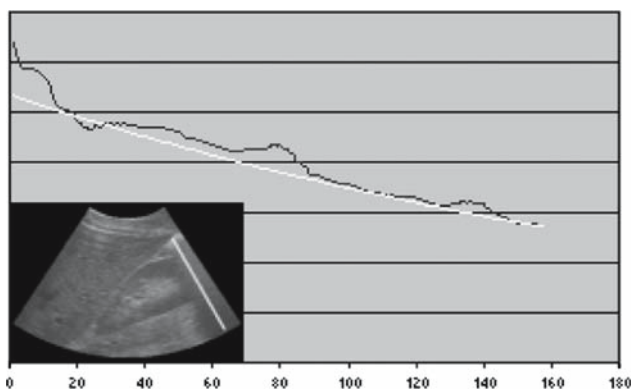


Fig. 7 Grey profile along the US scan line profile (*dark*) compared to the heuristic correlated function (*pale*); the corresponding direction indicated in the US image includes a shadow. X is the pixel position along the line, Y is the pixel intensity

grey-level values along a scan line. The definition of the heuristic function can be found in the appendix A.

A shadow is detected when the correlation coefficient is higher than a given threshold, and when the maximum acoustic interface is within a threshold distance to the skin. Figure 8 shows the result of this automatic method on a sample slice. Those thresholds have been selected from experiments on series of typical US images.

Effect of preprocessing on registration

Preprocessing effect upon registration was tested on a set of US slices and their corresponding CT oblique slice. Figures 9 and 10 present a typical example.

A systematic 2D registration test with an exhaustive exploration of the searching space showed that the cost function for the correlation ratio presents an obvious global minimum when the images are preprocessed (Fig. 10).

Similarity measure: correlation ratio

Definition

Let X be the base image (CT) and Y the floating image (US) to be matched to X . The CR is a functional similarity measure [6, 8, 11] based on conditional moments. It measures the functional relationship between the floating pixels and the base ones. Roche and colleagues [11] define the CR as:

$$\eta(Y|X) = 1 - \frac{1}{N\sigma^2} \sum_a N_a \sigma_a^2$$

where N is the number of overlapping pixels of X and Y for a given value of the unknown transform T , σ their variance in Y ; N_a is the number of overlapping pixels in X in which grey intensity value is a , σ_a their variance in Y . A more complete definition of the CR is given in the appendix B.

Motivation

Theoretically, the functional relationship between an intensity CT image and a gradient US image is not obvious;

Fig. 8 US shadow removal. *Left* image shows the shadow profiles induced by the ribs (*left*) and the colon (*right*). The generated mask is superimposed on right image

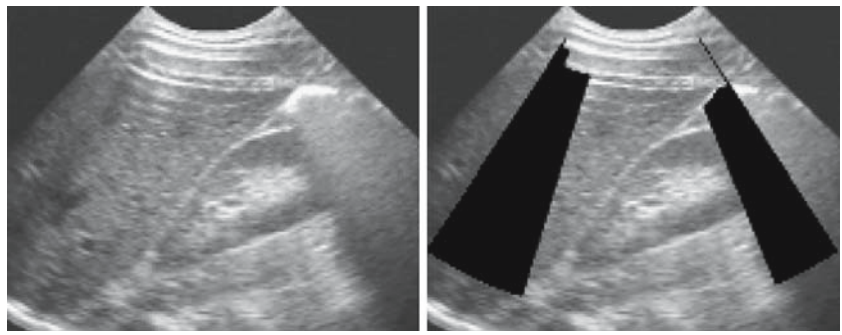


Fig. 9 A sample US slice and its corresponding CT oblique slice, after preprocessing

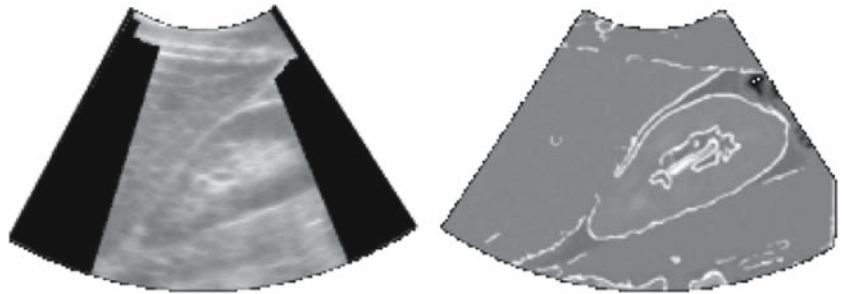
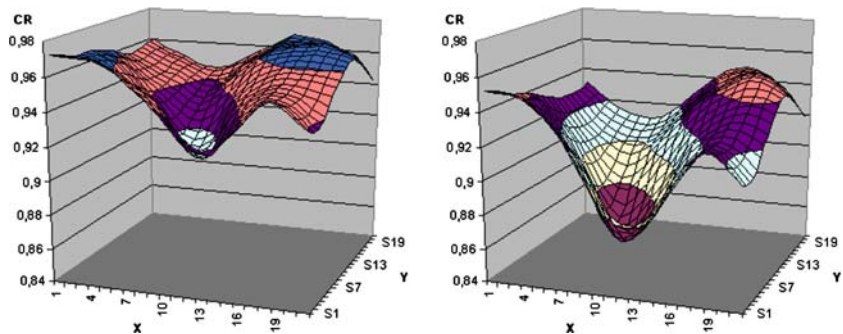


Fig. 10 Sample 2D registration of corresponding CT/US slices using CR, without preprocessing (*left*) and with preprocessing (*right*). The attraction basin is deeper in the latter case



thus a statistical similarity measure like normalized mutual information (NMI) should be more appropriate. However, although US images best enhance the boundaries between organs, they also convey information on the tissues, as each of them diffuses the US waves in its own way, in both intensity and frequency spaces. It was therefore assumed that CR is somehow justified, especially with the preprocessing steps presented earlier. We also chose CR against NMI for several reasons:

- CR looks smoother than NMI (Fig. 11); few local minima appear, even when sub-sampling; it is more accurate, discriminative and robust [8].
- Both CR and NMI algorithm complexities are $O(N)$, where N is the number of overlapping pixels in the ROI; however, CR requires less iterations, and is thus 30 to 60% faster.
- CR smooth profile is particularly adapted to PB search algorithm, because Brent minimization is based on the fitting of a parabola onto the function curve, as is described in the “Optimization scheme” section [19].

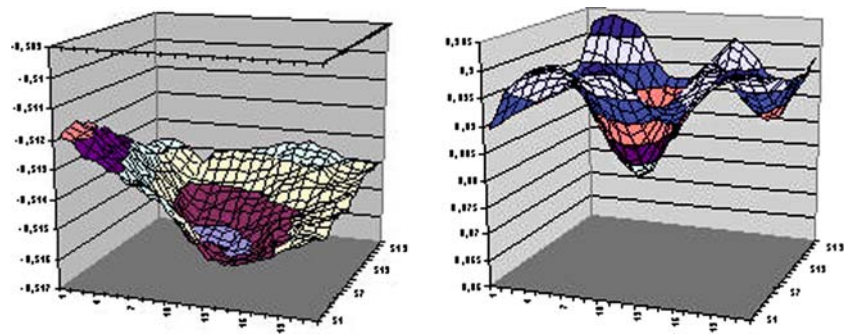
2.5D Correlation ratio

In the context of 2.5D US, the CR is computed as follows: for every rigid transform estimate, the CT volume is re-sliced in the plane of each US slice, using tri-linear interpolation. Then the region of interest (ROI) defined in each US image is superimposed onto the corresponding CT oblique slice. Finally, the CR is globally estimated on the set of pixels from the superimposed ROIs. In other words, the different US images are considered as a whole sparse set of pixels.

Optimization scheme: Powell–Brent

Powell–Brent is an iterative multi-directional search, based on the iterative fitting of a parabola to each 1D profile of the cost function in order to find 1D minima [4,6]. It appears as a fairly efficient search strategy when the differentiation of the cost function is unknown [4]. Our implementation is based on [19]. Some changes were applied since it was found that the method was too slow and too sensitive to local minima, which are frequent in image registration.

Fig. 11 NMI (left) and CR (right) profiles mapped along T_X and T_Y while registering two preprocessed images. Note the smooth aspect of CR and the deep unambiguous minimum



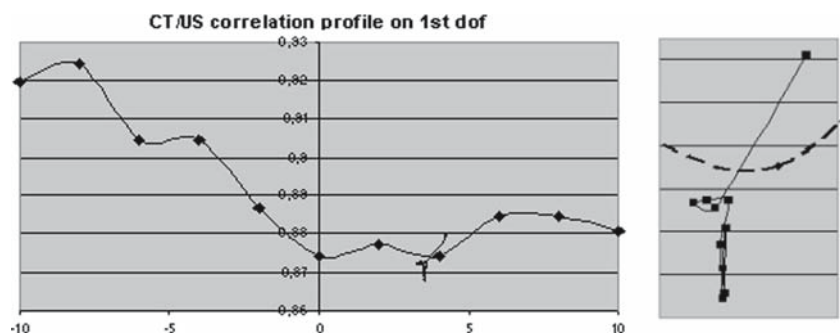
Initial attitude and parameter space

The minimization process consists in optimizing a 6D position and orientation vector. The initial attitude (IA) is defined by a point-to-point registration using the Arun algorithm [20]. Those points are anatomical landmarks visible on the US and CT data. The two poles (extreme points of the kidney along the CT Z axis) are identified and chosen on the CT images and on the US images. The other landmarks (at least one more) are generally chosen from a median transverse slice of the organ, for example at the connection with the ureter, or on the other side on the back of the kidney. On our data set the three or four points cited above were relatively well identifiable and repeatable. A reference system attached to the kidney is defined from those points: for instance the longitudinal axis of the CT kidney is defined as the Z axis of the parameter space. Finally, as Maes and colleagues suggests in [5], the better-conditioned parameter vector (T_X , T_Y , R_Z , T_Z , R_X , R_Y) is used for optimization.

Search interval limiting

In the original PB method, Brent 1D search interval for each DOF is defined by a time-consuming automatic bracketing. Based on repeated experiments about landmark selection for initial pre-registration and convergence from computed initial attitudes, an initial interval length of $2 \cdot \text{RMS}_{\text{ARUN}}$ (20 to 30 mm) around the IA was chosen.

Fig. 12 CR profile along the search interval for T_X . The converging curve represents Brent iterations around the global minimum of the profile



1D Initial search

Jenkinson et al. [6] proposed an initial systematic search on the rotational DOFs, previously to PB iterations, to make the solution “more reliable”. In the same idea, a two-stage approach was adopted with two step values. As PB method tends to converge to local minima, and as the similarity between registered CT and US shows in most cases a non-ambiguous global minimum (Fig. 11), we chose to perform an initial search before each Brent 1D optimization, with a step of 10% of the interval length, to find an approximate global minimum in that DOF (Fig. 12, left). Then Brent method is applied around that minimum in a very restricted interval (Fig. 12, right). With this two-stage approach, we observed that local minima were considerably avoided.

Search interval reduction

Since PB tends to reach the solution very fast (only the first iteration makes a significant step), it was decided, not only for computation time improvement but also to avoid peripheral local minima, to reduce the interval length at each Powell iteration with a factor of 1.5, until a certain limit is reached.

Normalizing search directions

After each Powell iteration, the new conjugate direction is normalized instead of being multiplied by the norm of the last found 6D minimum; indeed this minimum can be very small and can lead to early convergence in a local minimum.

Results

Materials and conditions

Several evaluations were performed on a single data set. This specific isotropic abdominal CT volume has $517 \times 517 \times 287$ voxels, the voxel size being 0.625 mm. Our registration experiments were carried out on a late acquisition CT, after iodine injection. Besides, for those experiments, 100 US images (axial and longitudinal) of the right kidney were acquired with a localized calibrated US probe, through anterior access. Those images whose pixel size is 0.3 mm are then resized to match CT voxel size in order to make further processing faster. In the final protocol, in order to stay in a framework compatible with a clinical use, we chose to carry out each registration on only five slices of various orientations out of the whole set. The ROI—parallelepipeds—are chosen to be tangent to the US kidney.

Measured data

The accuracy study consisted in performing two series of 10 registrations: on the one hand, by fixing the IA and changing the set of five US slices to be matched to the CT scan; on the other hand, by computing a new IA at each attempt, while keeping the same set of slices. Every IA is computed from the four anatomical points determined by the user on CT and US data at the axial and lateral extremities of the kidney.

Let us remind that a reference transform T_{BS} has been computed from the surface registration (cf. “Overview of the Method” section). The quality of a computed transform T is characterized by three measures (cf. Fig. 3):

- The distance between the two CT meshes obtained by transforming the CT data, respectively, by T_{BS} and T . This distance is computed between corresponding points of the initial mesh; mean and standard deviation values are presented for each registration trial.
- The distance between their centroids.
- The difference between their principal axes (angular distance).

Considering these three measures gives a better appreciation of the registration accuracy and success than a single RMS on registered data sets (CT and US).

Registration results

The registrations took on average $80s^2$ (from 11 to 170 s). The algorithm appeared to be stable because a second registration (i.e. another registration started from an IA that is the attitude provided by the first registration) kept the kidney with the same attitude.

² Tests performed on a Pentium IV 1.7 GHz.

Table 3 Matching statistics for varying set of US slices and constant IA (in mm and degrees)

Trial nb	Meshes distances (mean +/- std dev)	Centroid distances	Axis angular distances
1	5.6 (+/- 2.1)	2.7	4.1
2	6.8 (+/- 2.0)	3.2	6.7
3	5.3 (+/- 1.5)	3.2	1.3
4	6.2 (+/- 1.8)	4.0	5.3
5	5.8 (+/- 1.8)	2.3	9.5
6	6.0 (+/- 2.0)	4.3	4.7
7	4.1 (+/- 1.1)	2.7	5.5
8	5.1 (+/- 1.1)	3.4	5.2
9	5.1 (+/- 1.9)	3.0	5.4
10	3.7 (+/- 0.9)	2.1	2.9
Average	5.36 +/- 0.93	3.09 +/- 0.69	5.06 +/- 2.17

Table 4 Matching statistics for a constant set of US slices and varying IA (in mm and degrees)

Trial nb	Meshes distances (mean +/- std dev)	Centroid distances	Axis angular distances
11	3.7 (+/- 1.3)	1.3	4.3
12	3.2 (+/- 0.8)	2.9	1.5
13	8.8 (+/- 2.5)	5.7	10.5
14	5.7 (+/- 2.0)	4.0	5.3
15	6.7 (+/- 2.6)	5.1	8.2
16	6.0 (+/- 1.3)	5.8	3.2
17	11.9 (+/- 4.8)	8.1	6.4
18	6.5 (+/- 2.1)	5.0	7.6
19	6.2 (+/- 1.6)	5.6	6.0
20	5.4 (+/- 0.9)	4.1	4.3
Average	6.41 +/- 2.47	4.76 +/- 1.83	5.73 +/- 2.61

Tables 3 and 4 show matching statistics using the three measures listed above. On average, the IA computed from point-to-point registration sets the kidney at 7.7 ± 3.5 mm in translation and $9.2 \pm 4.0^\circ$ in inclination from the reference solution.

With 1 to 6 mm error in translation, and 1° to 7° in inclination, 15 registrations out of 20 have been considered successful. Figure 13 shows one of them. Trials 13 and 17 are considered as failures, whilst trials 5, 15 and 18 present higher errors on the kidney inclination than the average.

Discussion

We presented an original CT/US voxel-based registration method, dedicated to kidney surgery. The general approach, specific choices and results are discussed in this section.

The preprocessing of medical images before registration is sometimes put into question, because the cost function does

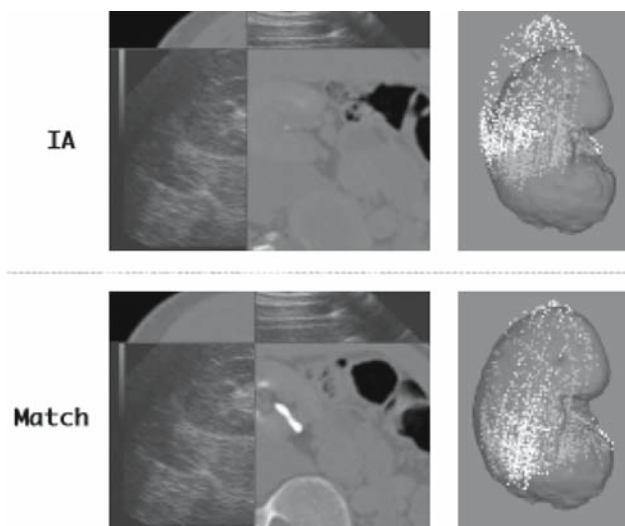


Fig. 13 Example of successful registration, showing superimposed CT/US slices and meshes. *Top* After determining the Initial Attitude using anatomical landmarks—*bottom* after registration using the intensity-based method

have a global minimum. By preprocessing data to be registered we chose to deepen the attraction basin of that minimum. The side effect is that local minima also gets deeper, even in a 20mm radius neighbourhood. However, by performing a systematic pre-search of the global minimum along every DOF, the risk of misregistration is decreased.

At any case, the removal of acoustic shadows appeared to be efficient in reducing the number of local minima. The method that was presented is fast and robust, provided that the quality of the images is good enough, in terms of contrast and regions homogeneity, particularly. This is an intrinsic drawback of using US data for which the image quality may be very variable from patient to patient but this drawback is the same when performing the puncture in the traditional way. Robustness to patient variability could probably be improved by integrating hybrid methods alternating registration and automatic segmentation stages such as presented in [21,22].

As regards the similarity measure, the CR seemed to be the best choice to make in the context of this application. Aside from the computation time question, the NMI as well as the other “statistic” and not “functional” measures presented a profile not sufficiently smooth to be used. Although the systematic search may bring closer to the large attraction basin of the global minimum, in the final iterations Brent’s algorithm is very likely to converge to one of the numerous surrounding local minima. Powell–Brent optimization scheme is much more successful with a smoother cost function such as CR.

Regarding the minimization process, a non-derivative method was chosen. Two well-known methods were good candidates: Powell–Brent and the Nelder–Mead simplex. The former was selected as it allows a better control of the convergence, of the search domain, etc. With a correct

implementation, the computation time is not greater for PB. However, authors like Maes et al. [5] have been able to use mutual information in derivative processes; some ideas may be taken out of it.

The developments were made, as well as several types of experiments, on one single data set from a young man; US data have been acquired through anterior access, which provides clear images. Some implementation choices were specific to this anterior access, particularly as regards the preprocessing of the images. Using the more traditional posterior access could make necessary the choice of some other preprocessing tools; however, the general approach can be kept identical.

Concerning accuracy, an average distance of 5.36mm to a bronze standard registration has been obtained where Penney and colleagues obtain an average 3.6mm. Whilst not totally satisfactory, those measurements which constrain the size of the target to be reached are reasonable in the context of this application (access to the pyelocaliceal cavities which size averages 1cm). This accuracy has to be confirmed on larger sets of data before envisioning the next stage involving needle guidance.

Conclusion

In this paper we addressed the problem of registering a CT-scan volume with a sparse set of localized US slices of the kidney. The intensity-based registration method is three-fold: a specific preprocessing of the images, a suited similarity measure and an efficient optimization process have been implemented. The preprocessing consists in decreasing the noise, enhancing the boundaries, and removing acoustic shadows. As for the similarity measure, correlation ratio was chosen against mutual information for its continuity and its computation time. The non-derivative Powell–Brent algorithm was implemented and allowed a precise control of the minimization process. Experimental results were compared to a “bronze standard” solution. Because of the theoretical difficulties in matching CT with US images, the presented results are satisfying in regard with accuracy and computation time. Nevertheless, we must bear in mind that their performances still depend on the choice of slices, the manual ROI, the manual IA and the search interval. Improving the robustness of the method with respect to those stages is a very important issue to focus on in the future.

Appendix A: Heuristic function used for shadow removal

Let $E = \{X_1, X_2, \dots, X_k, \dots, X_N\}$ be the profile of a scan-line in the US image. X_k is a grey level value along the profile. If i is the index of maximum grey level, let μ and σ be the

mean and standard deviation of $E' = \{X_i, X_{i+1}, \dots, X_N\}$. For k in $[i, N]$, let define $F = \{Y_i, Y_{i+1}, \dots, Y_N\}$ where $Y_k = f(k)$. The heuristic exponential function f is defined as:

$$f(k) = A \exp(ak); \quad A = \mu + \frac{3}{2}\sigma; \quad a = \frac{1}{N} \ln(X_N/A)$$

The correlation coefficient is:

$$\Gamma = \frac{\left| \frac{1}{N'} \sum_{k=i}^N X_k Y_k - m \right|}{\sqrt{\left| \left(\frac{1}{N'} \sum_{k=i}^N X_k^2 - m \right) \left(\frac{1}{N'} \sum_{k=i}^N Y_k^2 - m \right) \right|}}$$

$$m = \left(\frac{1}{N'} \sum_{k=i}^N X_k \right) \left(\frac{1}{N'} \sum_{k=i}^N Y_k \right)$$

Appendix B: Definition of the correlation ratio

Let X and Y be two sets of pixels/voxels. Ω is the overlapping set of pixels/voxels and N its cardinal. For a grey level a present in X , we note $\Omega_a = \{i | X(i) = a\}$ the set of pixel/voxel indexes in X having this grey level intensity and $N_a = \text{card}(\Omega_a)$

$$\text{Let } \sigma^2 = \frac{1}{N} \sum_{j \in \Omega} Y(j)^2 - m^2 \quad \text{and} \quad m^2 = \sum_{j \in \Omega} Y(j)$$

$$\sigma_a^2 = \frac{1}{N_a} \sum_{j \in \Omega_a} Y(j)^2 - m_a^2 \quad \text{and} \quad m_a^2 = \sum_{j \in \Omega_a} Y(j)$$

$$\text{The correlation ratio is } \eta(Y|X) = 1 - \frac{1}{N\sigma^2} \sum_a N_a \sigma_a^2$$

References

- Leroy A, Mozer P, Payan Y, Richard F, Chartier-Kastler E, Troccaz J (2002) Percutaneous renal puncture: requirements and preliminary results. *Surgetica'02*, pp 303–309
- Mozer P, Leroy A, Payan Y, Troccaz J, Chartier-Kastler E, Richard F (2005) Computer-assisted access to the kidney. *Med Robot Comput Assist Surg* 1(4): 58–66
- Wells WM, Viola P, Atsumi H, Nakajima S, Kikinis R (1996) Multi-modal volume registration by maximization of mutual information. *Med Image Anal* 1(1): 35–51
- Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P (1997) Multimodality image registration by maximization of mutual information. *IEEE Trans Med Imag* 16(2): 187–198
- Maes F, Vandermeulen D, Suetens P (1999) Comparative evaluation of multiresolution optimization strategies for multimodality registration by maximization of mutual information. *Med Image Anal* 3(4): 373–386
- Jenkinson M, Smith S (2001) A global optimization method for robust affine registration of brain images. *Med Image Anal* 5(2): 143–156
- Studholme C, Hill DLG, Hawkes DJ (1996) Automated 3-D registration of MR and CT images of the head. *Med Image Anal* 1(2): 163–175
- Sarrut D, Miguët S (1999) Similarity measures for image registration. European workshop on content-based multimedia indexing. IHMPT-IRIT, Toulouse, France, pp 263–270
- Pluim JPW, Maintz JBA, Viergever MA (2000) Image registration by maximization of combined mutual information and gradient information. *IEEE Trans Med Imag* 19(8): 809–814
- Haber E, Modersitzki J (2006) Intensity gradient based registration and fusion of multi-modal images, MICCAI 2006. LNCS 4191, pp 726–733
- Roche A, Malandain G, Pennec X, Ayache N (1998) The correlation ratio as a new similarity measure for multimodal image registration, MICCAI'98. pp 1115–1124
- Roche A, Pennec X, Malandain G, Ayache N (2001) Rigid registration of 3D ultrasound with MR images: a new approach combining intensity and gradient. *IEEE Trans Med Imag* 20(10): 1038–1049
- Penney GP, Blackall JM, Hamady MS, Sabharwal T, Adam A, Hawkes DJ (2004) Registration of freehand 3D ultrasound and magnetic resonance liver images. *Med Image Anal* 8(1): 81–91
- Jannin P, Fitzpatrick JM, Hawkes DJ, Pennec X, Shahidi R, Vannier MW (2002) Validation of medical image processing in image-guided therapy. *IEEE Trans Med Imag* 21(12): 1445–1449
- Langen K, Jones D (2001) Organ motion and its management, *Int J Radiat Oncol Biol Phys* 50(1):265–278
- Schwarz L, Richaud J, Buffat L, Touboul E, Schlienger M (1994) Kidney mobility during respiration. *Radiother Oncol* 32: 84–86
- Lango T (2000) Ultrasound guided surgery: image processing and navigation. PhD Thesis, Norwegian University of Science and Technology
- Czerwinski RN, Jones DL, O'Brien WD (1999) Detection of lines and boundaries in speckle images—application to medical ultrasound. *IEEE Trans Med Imag* 18(2): 126–136
- Press WH, Flannery BP, Teukolsky SA, Vetterling WT (1992) Numerical recipes in C: the art of scientific computing. Cambridge University Press, London
- Arun KS, Huang TS, et al (1987) Least-squares fitting of 2 3-D point sets. *IEEE Trans Pattern Anal Mach Intell* 9(5): 699–700
- Bansal R, Staiab L, Chen Z, Rangarajan A, Kniseky J, Nath R, Duncan J (1999) A minimax entropy registration framework for patient setup verification in radiotherapy. *Comput Aided Surg* 4(6): 287–304
- Ionescu G, Lavallée S, Demongeot J (1999) Automated registration of ultrasound with CT images: application to computer assisted prostate radiotherapy and orthopedics. In: Proceedings of MICCAI'99. LNCS, vol 1679, pp 768–777