



Automatic CT-ultrasound registration for diagnostic imaging and image-guided intervention

Wolfgang Wein^{a,*}, Shelby Brunke^b, Ali Khamene^a, Matthew R. Callstrom^c, Nassir Navab^d

^aImaging and Visualization Department, Siemens Corporate Research, Inc., 755 College Road East, Princeton, NJ 08540, USA

^bSiemens Medical Solutions, Inc., Ultrasound Division, Issaquah, WA, USA

^cDepartment of Diagnostic Radiology, Mayo Clinic, Rochester, MN, USA

^dComputer Aided Medical Procedures (CAMP), Technische Universität München, Germany

ARTICLE INFO

Article history:

Received 1 February 2008

Received in revised form 16 May 2008

Accepted 10 June 2008

Available online 19 June 2008

Keywords:

CT
Ultrasound
Registration
Fusion
Image-guided Intervention

ABSTRACT

The fusion of tracked ultrasound with CT has benefits for a variety of clinical applications, however extensive manual effort is usually required for correct registration. We developed new methods that allow one to simulate medical ultrasound from CT in real-time, reproducing the majority of ultrasonic imaging effects. They are combined with a robust similarity measure that assesses the correlation of a combination of signals extracted from CT with ultrasound, without knowing the influence of each signal. This serves as the foundation of a fully automatic registration, that aligns a 3D ultrasound sweep with the corresponding tomographic modality using a rigid or an affine transformation model, without any manual interaction.

These techniques were evaluated in a study involving 25 patients with indeterminate lesions in liver and kidney. The clinical setup, acquisition and registration workflow is described, along with the evaluation of the registration accuracy with respect to physician-defined Ground Truth. Our new algorithm correctly registers without any manual interaction in 76% of the cases, the average RMS TRE over multiple target lesions throughout the liver is 8.1 mm.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Conventional 2D ultrasound systems can be equipped with position sensors to perform 3D acquisitions of whole organs, and to obtain spatial information during a procedure. This is usually denoted as “3D Freehand Ultrasound” in the technical community (while ultrasound vendors often declare hand-swept imaging without tracking as “Freehand”). The fusion of such 3D ultrasound imaging with tomographic modalities can be used not only to improve navigation and ultrasound-based guidance for interventional procedures, but also to improve diagnostic value, e.g. for assessment of indeterminate lesions. This requires that the target anatomy is precisely registered between ultrasound and the pre-operative modality. Doing so in an automated manner is very challenging, and is the main focus of this work.

1.1. Clinical context

A common clinical problem is a patient with an indeterminate lesion, contained in the liver or kidney, identified by computed tomography (CT) or magnetic resonance imaging (MRI). Often, fur-

ther clinical work-up involves characterization of the lesion with the use of ultrasound (US) imaging. US imaging is utilized because this modality can often determine whether the lesion is likely benign or malignant due to the characteristic appearance of many lesions. However, it is sometimes difficult to correlate CT or MRI findings with US imaging due to inherent differences in the imaging methods. US imaging has several limiting factors that affect the use of this methodology for lesional identification and for its use as a guidance method. These factors include: (1) image acquisition is user dependent, (2) the field of view is limited, (3) US images are typically acquired off true axial, sagittal, or coronal planes with resultant difficulty correlating with these other cross-sectional imaging methods, (4) lesional identification can be difficult due to its echogenicity relative to the organ that is interrogated, and (5) the quality of the imaging is affected by the physical characteristics of the patient and overlying structures such as ribs, subcutaneous fat, and normal gas-containing structures. In this context, the fusion of CT and ultrasound can improve the diagnostic value to an extent beyond the “sum” of the individual modalities, potentially sparing an invasive biopsy, where a tissue sample for further pathological examination is obtained.

This leads to a second common clinical issue – the need to perform percutaneous needle biopsy or ablation of an indeterminate or malignant lesion contained in the liver or kidney. In the most

* Corresponding author. Tel.: +1 609 734 4477; fax: +1 609 734 3310.
E-mail address: wolfgang.wein@siemens.com (W. Wein).



Fig. 1. Radiofrequency ablation of a liver mass guided by ultrasound imaging in the CT suite.

common ablation technique, radio-frequency ablation (RFA), one or multiple ablation needles are inserted into the center of the lesion(s). Heat created around the needle tip causes coagulative necrosis of the tumor tissue (vanSonnenberg et al., 2005). Often, US image guidance is the method of choice for these procedures because it is performed real-time which allows precise placement of the biopsy needle or ablation device (Fig. 1). However, as with diagnostic imaging with US, the use of US for image-guidance can be difficult because the target lesion may not be visible or the imager may not be confident that a lesion identified with US imaging actually corresponds to the abnormality identified with CT or MRI (Wood et al., 2000). While many factors are important to achieve local tumor control of a targeted lesion when using percutaneous image-guided ablation devices, the most important one is accurate placement of the device in the center of the targeted tumor.

Many image-guided procedures are best performed with intermittent CT guidance, particularly when the target is relatively immobile in the retroperitoneum, lung, and bone. However, CT is of more limited usefulness as a guidance method for mobile targets, including the liver or kidney, due to the interrupted nature of the technique with superimposed patient breathing and motion. Multiple needle or device placement attempts may be necessary to successfully access a small target. This in turn increases the risk of possible complications such as bleeding.

One can clearly see that the use of tracked ultrasound, alongside with multi-modal registration of ultrasound to the pre-operative CT data, has a huge potential of improving ablation treatments. Many uncertainties regarding intra-operative tumor localization and delivery can be removed by precisely mapping the pre-operative and planning data into the coordinate system of the live ultrasound during the actual procedure (Crocetti et al., 2008).

1.2. Commercial systems

To the best of our knowledge, Ultraguide (est. 1996, Haifa, Israel) offered the first commercially available system that used magnetically tracked ultrasound transducers and biopsy instruments (Stippel et al., 2002). However, the clinical benefit of the technology over traditional fixed biopsy guides was never clearly demonstrated (Sheafor et al., 2000). The company went bankrupt in 2003.

The ultrasound vendor Biosound Esaote (Genoa, Italy) features a solution denoted *Virtual Navigator*, available on some of its products. It uses magnetic tracking as well, of both the ultrasound transducer and, optionally, an ablation electrode (Crocetti et al., 2008). The pre-operative CT data is integrated before the ablation procedure using manual registration.

Likewise, the *Real-time Virtual Sonography* system by Hitachi Medical (Tokyo, Japan) allows fusion of tracked ultrasound with CT or MRI, it has been applied for liver ablation and breast imaging (Arai et al., 2006).

Traxtal Inc. (Toronto, Canada) recently released the *PercuNav* system for Image-Guided Needle Interventions. It uses miniature electromagnetic tracking sensors integrated in the tip of ablation and biopsy needles, providing accurate navigation even for flexible instruments. Integration of pre-operative imaging and intra-operative tracked ultrasound is featured as well.

In the context of the described systems, automatic registration has the promise of improving the overall workflow by saving time and decreasing the human factor and subjectivity in aligning the pre- and intra-operative images.

1.3. Related work

A number of techniques for automatic registration of CT/MRI to ultrasound are described in the literature. In Roche et al. (2001), image-based registration of MRI to 3DUS is achieved by using both MRI intensity and gradient information in a similarity criterion based on correlation ratio (CR). Automatic registration on a single kidney CT/US data using CR as well, here by enhancing the CT intensities with major boundaries, is done in Leroy et al. (2004). In Penney et al. (2004), both MRI and US are remapped to an intermediate vessel probability representation using training data sets, then cross-correlation is used as similarity measure. In Wein et al. (2007b), a multi-component similarity measure involving weighted Mutual information is used on CT intensities and edge maps for rigid alignment with freehand ultrasound of the head and neck.

The mentioned methods all require manual initialization of the registration transformation, some need manual frame selection as well. In our work, we present a simulation of ultrasound from CT, which is realistic enough to allow for a stable registration, yet is computationally efficient at the same time. This has the added benefit that the simulation can be used by physicians or sonographers during planning, to get a feeling for the accessibility and optimal orientations even before the ultrasound exam, or the ultrasound-guided intervention. Furthermore, a novel similarity measure is developed, which is invariant to missing simulation details, yielding smooth properties and a global maximum at the correct alignment. We had presented a first version of it, along with preliminary clinical evaluation, in Wein et al. (2007a).

An ultrasound simulation based on a segmented CT scan, for training purposes rather than registration, has been proposed in Zhu et al. (2007).

2. Methods

2.1. Simulation of ultrasound from CT

In the following we explain the aspects of ultrasound physics relevant to our work. A more detailed introduction can be found in Zagzebski (1996). An ultrasound wave is partly reflected whenever a change in acoustic impedance is encountered in the imaged tissue. The acoustic impedance $Z = \rho c$ depends on the tissue density ρ and the speed of sound c . Ultrasound machines assume a constant $c = 1540$ m/s in human soft tissue, while a significantly different speed of sound occurs e.g. in air and bone. Table 1 lists the values for various tissue types, from Schneider et al. (1996). The ratio of an ultrasound wave intensity, reflected at a tissue interface with different acoustic impedances Z_1 and Z_2 is $(Z_2 - Z_1)^2 / (Z_2 + Z_1)^2$, given a specular interface with angle of incidence equal to the angle of reflection. The diffuse reflection Δr ,

Table 1

CT Hounsfield units μ and physical properties of human tissues (density ρ , speed of sound c , acoustic impedance Z)

Material	μ	ρ ($\frac{\text{kg}}{\text{m}^3}$)	c ($\frac{\text{m}}{\text{s}}$)	Z ($\frac{\text{kg}}{\text{mm}^2\text{s}}$)
Bone	1000	1912	4080	7.8
Muscle	10 ... 40	1080	1580	1.7
Liver	40 ... 60	1060	1550	1.64
Blood	40	1057	1575	1.62
Kidney	30	1038	1560	1.62
Brain	43 ... 46	994	1560	1.55
Water	0	1000	1480	1.48
Fat	-100 ... -50	952	1459	1.38
Air	-1000	1.2	330	0.0004

which is reflected straight back to the ultrasound transducer depends on the angle of incidence θ (see Fig. 2):

$$\Delta r(Z_1, Z_2, \theta) = (\cos \theta)^n \left(\frac{Z_2 - Z_1}{Z_2 + Z_1} \right)^2 \quad (1)$$

The exponent n describes the heterogeneity on the tissue interface, causing the amount of reflection to be more or less narrow around its perpendicular. We lack detailed physical knowledge from CT, hence we use $n = 1$, as it simplifies the equations and produces good results. Higher values would restrict the reflections of non-perpendicular interfaces, possibly missing to extract some features from the CT intensities. On the other hand, the similarity measure that will be explained later is to some extent capable of ignoring additional information not present in ultrasound. The transmitted intensity $t(Z_1, Z_2)$ does not depend on the angle of incidence, if refraction is neglected:

$$t(Z_1, Z_2) = 1 - \left(\frac{Z_2 - Z_1}{Z_2 + Z_1} \right)^2 = \frac{4Z_2Z_1}{(Z_2 + Z_1)^2} \quad (2)$$

The X-ray attenuation μ measured by a CT scanner is approximately proportional to the tissue density, see Fig. 3. Because tissue density is in turn proportional to acoustic impedance (as c is assumed constant), we can directly derive the incremental acoustic intensity reflection from it

$$\Delta r(\vec{x}, \vec{d}) = \left(\vec{d}^T \frac{\nabla \mu(\vec{x})}{|\nabla \mu(\vec{x})|} \right)^n \left(\frac{|\nabla \mu(\vec{x})|}{2\mu(\vec{x})} \right)^2 \quad (3)$$

$$\text{for } n = 1: \Delta r(\vec{x}, \vec{d}) = \left(\vec{d}^T \nabla \mu(\vec{x}) \right) \frac{|\nabla \mu(\vec{x})|}{(2\mu(\vec{x}))^2} \quad (4)$$

$$t(\vec{x}) = 1 - \left(\frac{|\nabla \mu(\vec{x})|}{2\mu(\vec{x})} \right)^2 \quad (5)$$

where $\mu(\vec{x})$ is the CT attenuation value at position \vec{x} , $\nabla \mu(\vec{x})$ its spatial derivative, and \vec{d} a unit vector denoting the direction of the ultrasound wave propagation. The scalar multiplication of \vec{d} with the normed CT gradient vector yields the angular dependency equivalent to $\cos(\theta)$. The ultrasound wave intensity is reduced

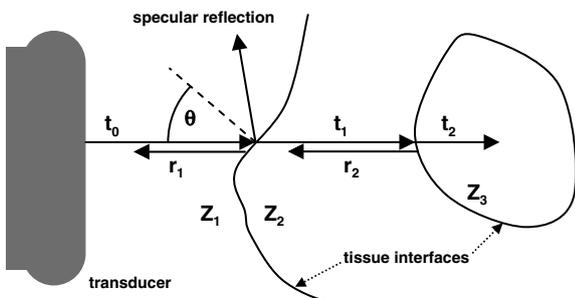


Fig. 2. Principle of ultrasonic transmission and reflection at multiple tissue interfaces.

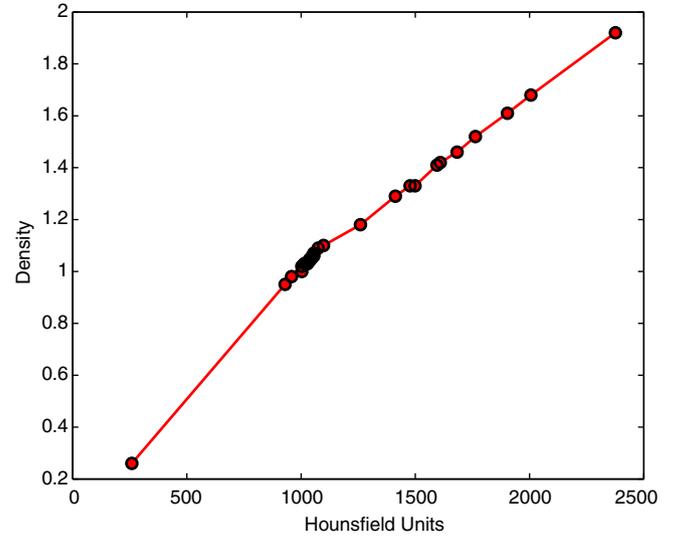


Fig. 3. Plot of CT Hounsfield units against tissue density, values from Schneider et al. (1996).

according to $t(\vec{x})$ at each tissue interface, while $\Delta r(\vec{x}, \vec{d})$ contributes to the wave intensity detected by the probe. Integrating over this reflection and transmission behavior along an ultrasonic scanline yields:

$$I(\vec{x}) = I_0 \exp \left(- \int_0^{\lambda_x} \left(\frac{|\nabla \mu(\vec{x}_0 + \lambda \vec{d})|}{2\mu(\vec{x}_0 + \lambda \vec{d})} \right)^2 d\lambda \right) \left(\vec{d}^T \nabla \mu(\vec{x}) \right) \frac{|\nabla \mu(\vec{x})|}{(2\mu(\vec{x}))^2} \quad (6)$$

where I_0 is the original intensity of the ultrasound pulse, we define it as $I_0 = 1$. In addition, we apply a log-compression, which amplifies smaller reflections. Its parameter a is visually determined, and resembles the dynamic range knob on the ultrasound machine. This yields the resulting value of the simulation:

$$r(\vec{x}) = \frac{\log(1 + aI(\vec{x}))}{\log(1 + a)} \quad (7)$$

For a linear array probe, the integral in Eq. 6 can be computed efficiently by traversing the columns in the simulated ultrasound image from top to bottom while updating the transmitted intensity based on the interpolated CT intensity and gradient values. For curvilinear arrays, we compute the image row-wise from top to bottom, while using an auxiliary channel storing the remaining transmitted ultrasound wave intensity (starting with 1 in the first row). For every pixel, this transmission value is retrieved by linear interpolation from two pixels in the row above, according to the ultrasound ray angle derived from the curvilinear geometry.

This provides a means to simulate large-scale ultrasonic reflection at tissue boundaries, and the related shadowing effects at strong interfaces like bone. However, individual tissue types have specific echogeneity and speckle patterns by themselves, based on the microscopic tissue inhomogenities. There is no simple relationship between tissue echogeneity and CT hounsfield units, therefore we add an intensity mapping $p(\mu(\vec{x}))$ on a narrow soft-tissue range to the simulated large-scale reflection $r(\vec{x})$. We use a simple polynomial function, based on a number of correspondences (liver tissue, liver vasculature, kidney, gall bladder) between CT/CTA intensities and tissue echogeneity in ultrasound, see Fig. 4. Fig. 5 depicts the simulation result for a transverse liver image, computed from a native CT scan.

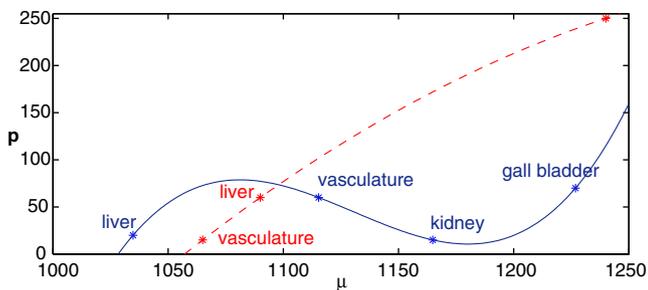


Fig. 4. Intensity mapping p for CT (dashed) and portal-venous CTA (solid) foot tissue. Note that the liver-vasculature relation is inverted in the two modalities.

2.2. Registration algorithm

2.2.1. Automatic frame selection

Since we simulate ultrasound imaging effects with respect to the probe geometry, the original B-mode scan planes of the sweep have to be used rather than a 3D reconstruction. Neighboring frames of the freehand sweep contain similar information, hence we use always the one out of n frames that has the highest image entropy. This assures that frames which contain unique fine vascularity, that can be located in CT as well, are picked for registration. If two neighboring frames have the highest entropy out of their group of n , only one of them (again with the highest entropy) is used. Furthermore, a threshold is used to discard frames at the beginning and end of the sweep with little structures. In our experiments, n was defined to yield 20–30 frames per sweep for registration.

2.2.2. Similarity measure considerations

It seems appropriate to use statistical similarity metrics like mutual information (MI) and correlation ratio (CR) for assessing the correspondence of original CT and ultrasound intensities. In their general formulation, however, they do not work well for our registration problem, since there are too many possible configurations where the joint entropy is minimal (for MI), or the intensities from one image can be predicted well from the other one (for CR). At correct alignment of CT and ultrasound, they typically produce only a small local optimum. Known approaches for restricting the possible intensity distributions are distance metrics to joint histograms learnt from correct registrations (see e.g. Guetter et al. (2005), Kullback–Leibler divergence), as well as bootstrapping parameters for a polynomial intensity mapping in the actual registration process itself (Roche et al., 2001). In both cases, very important information is disregarded, as e.g. small vascularity is essential for a correct registration within the liver, but due its

appearance on a relatively small fraction of the image content, it would neither affect a joint histogram or a least-squares estimate of a polynomial intensity mapping. Since CT attenuation measurements are mostly reproducible, we will use the constant mapping function p defined in Section 2.1 in conjunction with a linear model.

2.2.3. The LC^2 similarity measure

In a correlation ratio framework, the parameters of the registration transformation T are modified in order to maximize

$$CR = 1 - \frac{\sum_{x \in \Omega} (U(x) - f(\mu(T(x))))^2}{|\Omega| \text{Var}(U)} \quad (8)$$

with f denoting the mapping function which estimates the intensities of the US image U from the transformed CT image μ , and Ω the shared image domain. If a linear mapping $f(\mu) = \alpha\mu + \beta$ is assumed, Eq. (8) can be directly related to the common normalized cross-correlation (NCC) similarity metric, see Roche et al. (1999) for the derivation.

For a pixel intensity in the US image, the relative contributions of large-scale reflections and general tissue echogeneity are unknown. Hence both the mapped CT intensity $p(\mu)$ and the simulated reflection r have to be integrated in a correlation framework with the US intensity. Using the notation $p_i = p(\mu(T(\vec{x}_i)))$, $r_i = r(T(\vec{x}_i))$, $u_i = U(\vec{x}_i)$ for the intensity triple at a certain voxel, we define the intensity function as

$$f(\vec{x}_i) = \alpha p_i + \beta r_i + \gamma \quad (9)$$

The unknown parameters α , β and γ then have to minimize

$$\left\| M \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} - \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \right\|^2; \text{ where } M = \begin{pmatrix} p_1 & r_1 & 1 \\ \vdots & \vdots & \vdots \\ p_n & r_n & 1 \end{pmatrix} \quad (10)$$

Therefore the solution is

$$\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = (M^T M)^{-1} M^T \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} \sum p_i^2 & \sum p_i r_i & \sum p_i \\ \sum p_i r_i & \sum r_i^2 & \sum r_i \\ \sum p_i & \sum r_i & n \end{pmatrix}^{-1} \begin{pmatrix} \sum p_i u_i \\ \sum r_i u_i \\ \sum u_i \end{pmatrix} \quad (11)$$

Direct inversion of the symmetric matrix $M^T M$ results in a closed-form solution for the parameters. They are then inserted in Eq. (8) to yield a novel registration similarity metric, which we denote *Linear Correlation of Linear Combination* (LC^2). It assesses the correlation of US intensities u_i and a linear combination with unknown weights of signals p_i, r_i extracted from CT. The value of LC^2 is constant with respect to brightness and contrast changes of the US image (as NCC), but also independent to how much of the two described physical effects contributes to the image intensities. The

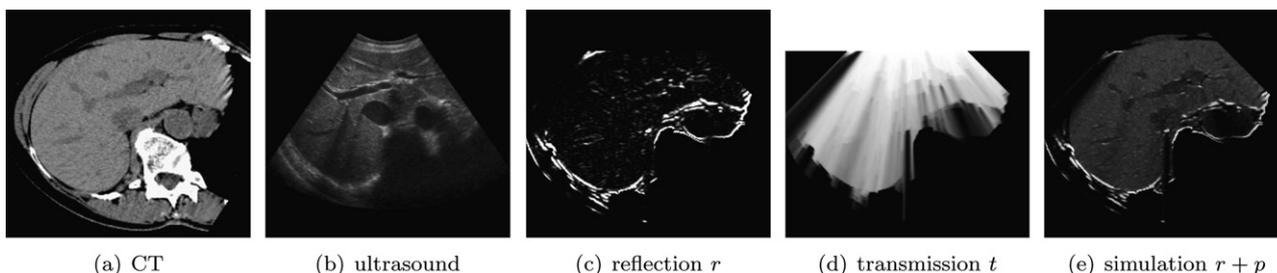


Fig. 5. Simulation of ultrasonic effects from CT. The original images are depicted in (a) and (b). The simulated large-scale reflection is shown in (c), the auxiliary transmission image in (d). The final simulation outcome is (e), where reflection and echogeneity estimated from CT are combined. (a) CT, (b) ultrasound, (c) reflection r , (d) transmission t and (e) simulation $r + p$.

latter is important, since e.g. hepatic vasculature or the gall bladder is represented mostly by p (different intensities due to echogeneity in ultrasound, no borders), while large-scale tissue interfaces correspond to r (strong edge in ultrasound, comparable intensities on both sides).

This implicit computation of the parameters α , β and γ during every pose evaluation equals a simultaneous optimization of radiometric and geometrical registration parameters (with the radiometric parameters providing information for simulating US images from the CT volume). Fig. 6 illustrates this by showing the simulated intensity according to $\alpha p_i + \beta r_i + \gamma$ for an aligned and displaced image of the liver. For the aligned image, the parameters α and β are higher, denoting a good least-squares matching of the structures. For the displaced image, they are very small, causing both the vasculature and liver border to almost disappear. The value of γ is expressed in a normalized intensity range $[0 \dots 1]$.

Note: The fact that the weights for p and r are unknown, suggests that a higher-dimensional mutual information (HMI) approach could be used as well (Gan and Chung, 2005). Here, each p, r and u would represent one axis in a three-dimensional joint probability distribution. We had investigated this approach, computing modified versions of MI and CR on a 3D joint histogram; it however resulted in an unstable similarity metric (see also Section 3.3). One reason is that the reflection term r has an intensity distribution containing mostly small values (no reflections) and few yet important large values (representing reflections at tissue interfaces). A non-linear histogram equalization approach would be a pre-requisite to use r in such a framework. Another problem, as pointed out before in Section 2.2.2, is the unconstrained huge number of intensity configurations that lead to a local optimum.

Eqs. (11) and (8) can be computed once for a set of US-CT image pairs, individually for every frame in the set (as we did in Wein et al., 2007a), or locally for arbitrary image regions (see Section 2.2.4 below). In all but the first case, the mean of the resulting LC^2 values constitutes the cost function for optimization.

For disregarding fine speckle information in the registration (whose correspondence we cannot extract from CT), and speed-up of the computation, the US images are down-scaled to $\sim 128 \times 100$ pixels (by averaging an integer number of pixels in each dimension). The top 3 cm along the ultrasonic rays are ignored for the measure computation, since they contain only com-

pressed subcutaneous tissue (which we visually confirmed on a large number of data sets).

2.2.4. Local LC^2 computation

Generally, it can not be assumed that Eq. (9) is valid over the whole content of a registered image pair. One reason, as pointed out above, is that different anatomic structures yield different weightings of the echogeneity and reflection terms p and r . Apart from that, every ultrasound machine features settings that locally optimize the image quality, such as the time gain compensation (TGC) curve or the number and location of focal zones. Last but not least, orientation-dependant artifacts result in upper structures influencing the imaging of anatomy further away from the probe, which might not always be accurately reproduced by our simulation from CT.

We therefore compute the LC^2 similarity measure for local patches centered around every pixel in each image pair. This extension is similar to the local normalized cross-correlation (LNCC) similarity metric, which has particularly proven useful for 2D–3D registration (Khamene et al., 2006). The patch size has to be chosen correctly; if it is too small, Eq. (9) will always hold, therefore not decreasing the image similarity with larger misalignment. If it is too large, it does not hold for the correct alignment, since structures reflecting different weightings for p and r share one patch. In a robustness study (see Section 3.3) carried out for different sizes, we obtained 11×11 pixels as the optimal patch size.

As a further advantage of the local LC^2 computation, we can drop the intensity mapping for contrasted CT scans (Fig. 4). The resulting inverse relationship between CT and US intensities of vasculature causes a negative α value for the respective patches (which would not be valid for the remainder of the image). This increases the local accuracy particularly of small vascular structures, as some ambiguity can be introduced by the mapping – one can see in Fig. 4 that up to three CT intensities are mapped to the same US intensity.

2.2.5. Optimization strategy

An initial estimate of the orientation is obtained from the tracking setup, if the patient is supine. Otherwise, an approximate angle around the cranio-caudal axis is entered manually. The large-scale translation is determined by performing an exhaustive search of the translation space. Here, a skin surface clamping approach is used, skipping physically impossible transducer locations, as in Wein et al. (2007b). All configurations within the evaluated 3D-grid suggesting an optimum, are further locally optimized with respect to the translation. From the configuration which in turn yields the best result, all six parameters of the rigid transformation are refined. As an optional last step, an optimization is executed on all rigid and three selected affine transformation parameters (henceforth denoted *semi-affine*). These are the two scaling parameters and the one shearing of the sagittal plane, since respiratory motion mainly causes deformation in that plane (Rohlfing and C.R. Maurer, 2004). For all local optimizations, the Amoeba Simplex algorithm is used, as described in Press et al. (1992), chapter 10.

3. Experiments

3.1. Setup

In order to evaluate the performance of the registration algorithm, a study on abdominal data of 25 patients was performed. Patients were included that had one or more indeterminate lesions contained in either the liver or kidney, that measured ≥ 0.5 cm and ≤ 5 cm in diameter, diagnosed on a prior contrast-enhanced

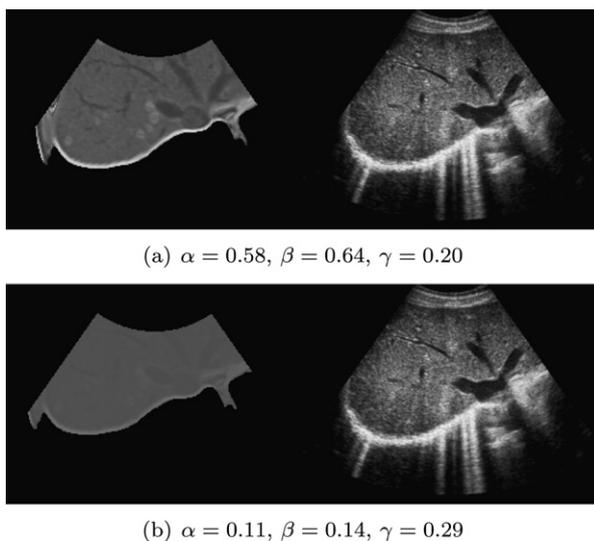


Fig. 6. The effect of simultaneous simulation and registration. The left column shows the simulation from CT using the parameters resulting from the LC^2 computation. (a) is well registered, (b) is 1 cm displaced.

CT examination. Patients with cardiac pacemakers or defibrillators were excluded from the study. No patients were excluded on the basis of body habitus.

For focal liver lesions, a biphasic contrast-enhanced examination of the abdomen (delay of 45 and 70 s) was performed. For focal renal lesions, a triphasic contrast-enhanced examination (delay of 45, 70 and 180 s) was done. In either case, image reconstruction slice thickness of ≤ 2 mm was required.

Since the study is ultimately targeted towards the interventional application, the decision was made to use a magnetic tracking system. The small tracking sensor can be attached right to the ultrasound transducer (optionally inside a sterile plastic wrap), and does not require a line of sight to the transmitter. This is advantageous in the RFA setup with limited space (see Fig. 1), and allows tracking of needles as well. Our freehand ultrasound system uses a 3D Guidance tracking system (Ascension Technology Corp., Burlington, VT USA) with the flat transmitter option, which slides below the patient mattress and avoids tracking errors induced by metal in the patient bed. Images from a Siemens Sequoia ultrasound machine are fed via progressive RGBS video into a PC with frame grabber. The position sensor was affixed to the transducer using hot-melt adhesive, and the method presented in Wein and Khamene (2008) was used to compute the spatial and temporal calibration.

Transverse liver sweeps on 21 patients (see e.g. Fig. 7) and four kidney sweeps, acquired during breath-hold on inspiration, were used. They were co-registered with the portal-venous phase of the CT scans, due to the conspicuity of the portal and hepatic veins in this phase of imaging. For optimal visualization, the ultrasound exam was executed in various setups - 14x supine, 7x left posterior oblique (LPO), 1x right posterior oblique (RPO), 3x decubitus.

3.2. Registration results

After manually aligning each of the data sets, 5–9 point correspondences on anatomical landmarks, mostly vessel bifurcations in the liver, were selected by an expert. In order to define those landmarks truly in 3D, we visualized both original ultrasound frames and an arbitrary number of cross-sections, compounded using the direct MPR technique presented in Wein et al. (2006), each with the respective CT plane (see Fig. 7b). Using a linked poin-

ter and superimposition options, the physician could precisely locate vessel bifurcations.

While the obtained point correspondences serve as *fiducial* landmarks, additional *target* point correspondences were defined, depicting the lesions to be ablated. Typically, the center of multiple small lesions, or distinguished spots in larger lesions close to peripheral vasculature, were used, creating 1–5 point correspondences (in one atypical case 10), which indicate where the highest registration accuracy is desired for optimal treatment. The point localization errors are expected to be higher than for the fiducials.

For registration based on point correspondences, the rigid motion between ultrasound and CT was computed according to Walker et al. (1991). Table 2 lists the mean, minimum and maximum of the root-mean-square (RMS) residual distances of the fiducial and target points (henceforth denoted as fiducial registration error FRE, target registration error TRE) for all data sets. They have been computed for point-based registration based on the fiducials alone, and both fiducial/target points; as well as rigid & semi-affine automatic registration using our methods.

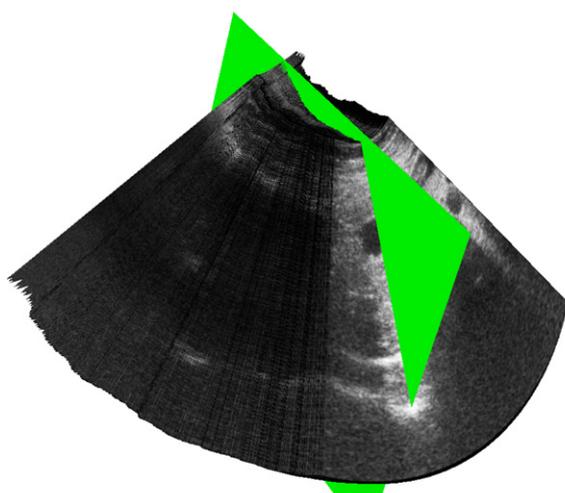
The automatic registration completes for 19 patients with an average computation time of ~ 28 s (C++ implementation executed on an Intel Core 2 Duo 2.2 GHz notebook). Two difficult supine data sets, as well as three LPO and one decubitus data, had to be roughly manually aligned in order for the automatic algorithm to converge. At the initial estimate, before the translation search, the FRE was between 13 and 71 mm (the flat transmitter of the positioning system was placed similarly below each patient).

In 59% of the cases, automatic affine registration yielded a lower TRE than point-based registration of the fiducials (55% for rigid).

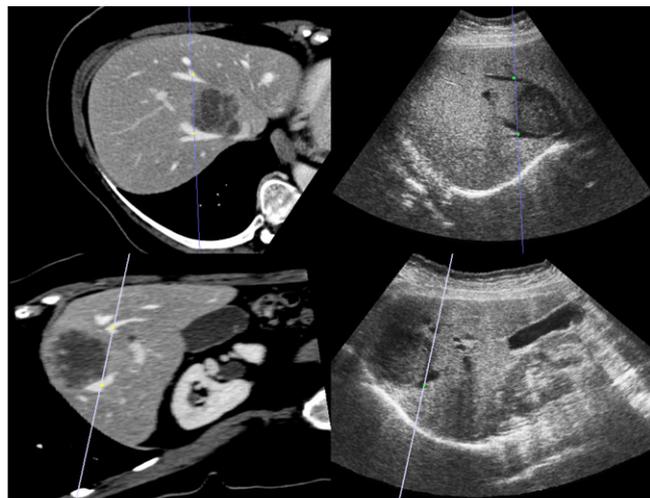
Table 2

Average registration accuracy in mm for 25 patients, expressed as root-mean-square fiducial registration error (FRE) and target registration error (TRE)

	Mean		Min		Max	
	FRE	TRE	FRE	TRE	FRE	TRE
Point-based (fiducial)	5.0	9.7	2.3	2.8	11.5	28.4
Point-based (fid. + target)	5.8	5.4	2.4	2.2	11.9	11.1
Automatic rigid	10.4	9.0	4.5	3.0	18.7	22.1
Automatic affine	9.5	8.1	2.8	3.0	15.3	21.5



(a) transverse liver sweep



(b) CT MPR planes alongside ultrasound frame (top) and reconstruction (bottom)

Fig. 7. Typical liver data, from a patient with a large metastasis near the hepatic vein confluence. The green plane in (a) corresponds to the ultrasound reconstruction in the lower right. (a) transverse liver sweep and (b) CT MPR planes alongside ultrasound frame (top) and reconstruction (bottom).

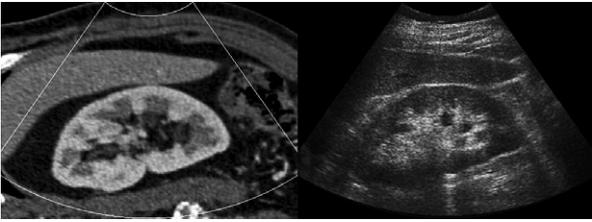
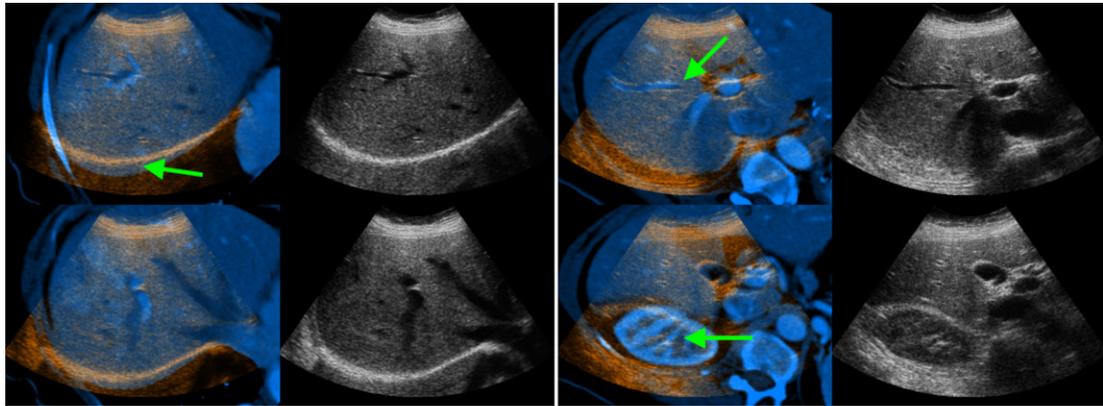
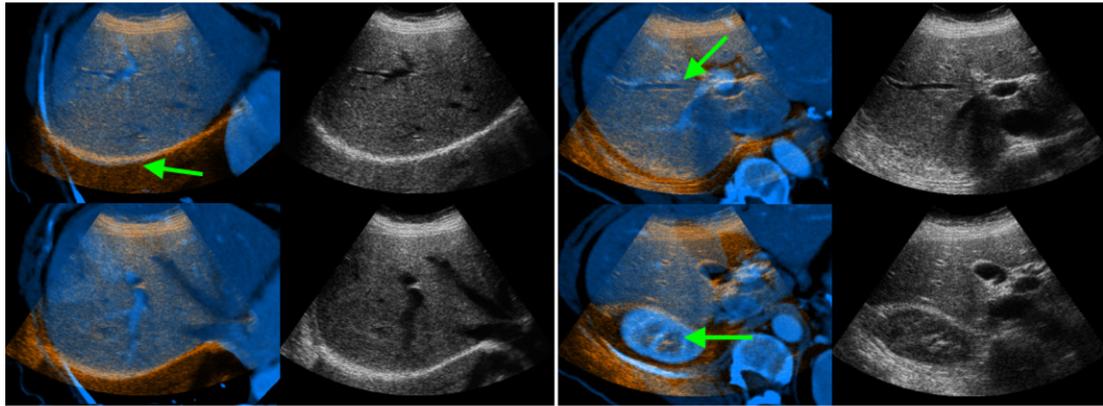


Fig. 8. Longitudinal image of a right kidney, aligned by automatic rigid registration.

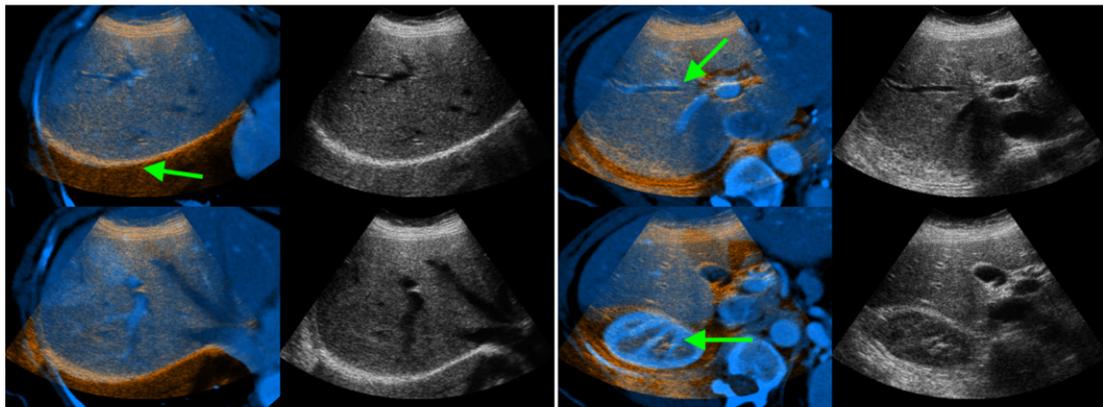
This can be attributed to the fact that the image-based technique incorporates information throughout the 3D sweep, trying to find an alignment of all structures visible in both modalities. On the other hand, even careful definition of unique fiducial landmarks does not necessarily guarantee a small TRE, particularly if the clinical targets are not in their vicinity. The FRE after point-based registration of the fiducials can obviously not be lowered by automatic registration, since it represents their residual error. It is composed mostly of inaccuracy in the point selection, tissue deformation between the two modalities, tracking and calibration errors. In 73% of



(a) registration based on fiducial point correspondences, FRE = 6.4mm



(b) automatic rigid registration, FRE = 9.8mm



(c) automatic affine registration, FRE = 8.1mm

Fig. 9. Comparison of point-based, rigid and affine registration. The patient was positioned left posterior oblique (LPO), the Sequoia Clarify option was enabled, enhancing vasculature. The 1st and 3rd column depict a color overlay of CT and US, 2nd and 4th column show the original US (for better visualization, the reader is referred to the web version of this article). Arrows point to anatomical clues that visualize the quality of alignment. (a) registration based on fiducial point correspondences, FRE = 6.4 mm (b) automatic rigid registration, FRE = 9.8 mm and (c) automatic affine registration, FRE = 8.1 mm.

the cases, the TRE after affine registration was lower than after rigid registration. This suggests that the semi-affine model can eliminate most of the large-scale deformations, particularly induced by CT and ultrasound exams in different respiratory configurations. As expected, this TRE difference was especially high for non-supine patient positioning during ultrasound. For the kidney cases, the alignment after automatic rigid registration looks visually excellent (see Fig. 8), although it was sometimes difficult to precisely locate fiducials. Here, the semi-affine model provided no improvement over the rigid one.

Fig. 9 illustrates the results for a liver sweep in LPO setup. Matching the fiducials results in well aligned vasculature, however gross structures such as the lower liver surface are misplaced. Our automatic method with a rigid model in turn correctly lines up the liver surface, while the matching of the vasculature and kidney is worse. Only the semi-affine model accurately aligns the liver border, kidney and features within the liver gland.

3.3. Similarity measure comparison

In order to evaluate the LC^2 metric, we performed a randomized robustness study on a typical liver sweep. From the fiducial ground truth, the translation parameters were individually displaced up to ± 20 mm and a local registration was executed, repeated 100 times. The box-and-whisker plot in Fig. 10 depicts the resulting distribution of the FRE error values. Our LC^2 measure with global, per-frame and local parameter estimation (see Section 2.2.3) was compared against MI, CR and higher-dimensional MI (HMI). For MI and CR, the sum of simulated echogeneity and reflection $p+r$, as shown in Fig. 5e, was inserted into a joint histogram against the US intensities u . For HMI, each p , r and u were used individually in a 3D joint histogram.

For each of the similarity measures, outliers (i.e. failed registrations) were defined by separating two clusters in a 2D-plot of the FRE against the similarity measure value (with the cluster of outliers having low similarity & high FRE values). The maximum initial displacement for the successful registration runs then defined the capture range of the measure. As can be seen in Table 3, global LC^2 has the highest capture range, and local LC^2 the highest accuracy (underlined values). Higher-dimensional MI fails for all trials.

Based on these results, we decided to solely use the local LC^2 measure due to its unmatched accuracy; the grid size of the global optimization (Section 2.2.5) was adjusted according to its capture range. In similar, systematic randomized studies, further algorithm

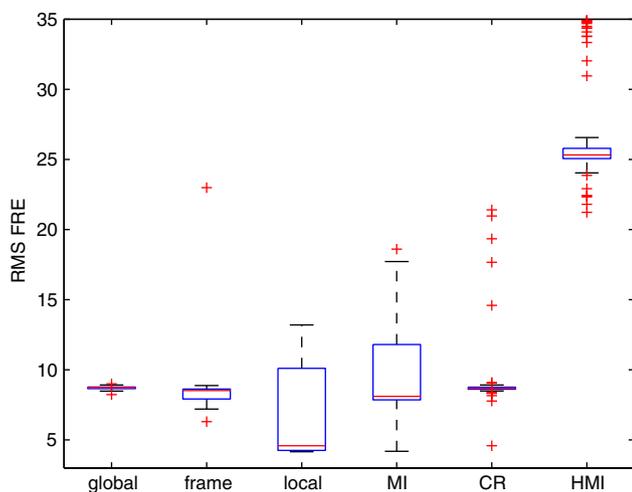


Fig. 10. Robustness of the different LC^2 variants and statistical similarity measures.

Table 3

Outliers, FRE and capture range of different measures

	LC^2			Statistical		
	Global	Frame	Local	MI	CR	HMI
Outliers (%)	0	1	35	29	7	100
Median FRE (mm)	8.7	8.5	4.6	8.1	8.7	25.3
Capture range (mm)	15.8	13.7	6.7	3.8	10.1	0

parameters were optimized with respect to the overall accuracy and robustness, such as the local LC^2 patch size, number of images per sweep used for registration, as well as the image resolution.

4. Conclusion

We have presented novel methods for simulation of ultrasonic effects from CT data, as well as the new LC^2 similarity metric that is robust with respect to unknown parameters in the simulation. Together, they allowed us to develop a new, fully automatic image-based algorithm for registering 3D freehand ultrasound sweeps with CT, each acquired in breath-hold.

The methods have been evaluated in a diagnostic fusion study on 25 patients with indeterminate lesions of the liver and the kidney. We established Ground Truth registration by carefully defining fiducial and target point correspondences in all data sets. The average RMS TRE of our automatic method with a semi-affine transformation model is 8.1 mm, better than point-based registration of the fiducial landmarks (9.7 mm). Point-based registration of both fiducial and target points yielded an average TRE of 5.4 mm. The automatic registration succeeded without manual pre-alignment in 76% of the cases, with an execution time of less than 40 s. This stands against a tedious definition of point correspondences, which took us ~ 10 min per patient in this study. An average RMS TRE value over multiple lesions of 8.1 mm is acceptable for overall organ registration, particularly since it represents an improvement upon the point-based registration of vascular features. We expect our method to greatly increase the acceptance of multimodal fusion for diagnosis and treatment, since it provides a simple workflow and enables more precise registration.

Using a global semi-affine model, the algorithm is at this point not able to fully compensate organ deformations, occurring mainly due to different patient setup or inhalation in each of the imaging modalities (the RMS TRE for supine data sets is 6.2 mm, and 10.3 mm for the remainder). We are currently investigating approaches to expand our methods towards deformable transformation models. Paired with the robust local LC^2 metric, the level of detail of our ultrasound simulation is sufficient for global alignment; however a more complex simulation model (including e.g. refraction and multiple reflections) might be required for non-linear registration. For precise interventional navigation, the TRE on single lesions has to be lowered significantly, weighting of the local similarity metric with respect to the targets is an additional option to achieve this. The global initialization of difficult cases, i.e. where ultrasound imaging characteristics are challenging, is another current limitation. Last, but not least, strategies to provide real-time compensation of respiratory motion with image-based techniques is left for future work.

Acknowledgement

We would like to thank the Drs. J.W. Charboneau, T.D. Atwell and M.A. Farrell from the Diagnostic Radiology Department at Mayo Clinic for their support of the study.

References

- Arai, O., Shinomura, R., Mitake, T., Sawaki, A., Satake, H., 2006. Real-time virtual sonography (RVS) for breast imaging. *Ultrasound in Medicine & Biology* 32 (Supplement 1), P107–P108.
- Crocetti, L., Lencioni, R., DeBeni, S., See, T., Pina, C., Bartolozzi, C., 2008. Targeting liver lesions for radiofrequency ablation: an experimental feasibility study using a CT-US fusion imaging system. *Investigative Radiology* 43, 33–39.
- Gan, R., Chung, A.C., 2005. Multi-dimensional mutual information based robust image registration using maximum distance-gradient-magnitude. In: *IPMI 2005 Proceedings*, vol. 3565 of *Lecture Notes in Computer Science*, pp. 210–221.
- Guetter, C., Xu, C., Sauer, F., Hornegger, J., 2005. Learning based non-rigid multimodal image registration using kullback–Leibler divergence. In: *MICCAI 2005 proceedings*, vol. 3750 of *Lecture Notes in Computer Science*, pp. 255–262.
- Khamene, A., Bloch, P., Wein, W., Sauer, F., Svatos, M., 2006. Automatic portal image based patient positioning for radiation therapy. *Medical Image Analysis* 10, 96–112.
- Leroy, A., Mozer, P., Payan, Y., Troccaz, J., 2004. Rigid registration of freehand 3D ultrasound and CT-Scan kidney images. In: *MICCAI Proceedings*, p. 837ff.
- Penney, G., Blackall, J., Hamady, M., Sabharwal, T., Adam, A., Hawkes, D., 2004. Registration of freehand 3D ultrasound and magnetic resonance liver images. *Medical Image Analysis* 8, 81–91.
- Press, W., abd, W.T., Vetterling, S.T., Flannery, B., 1992. *Numerical Recipes in C*, second ed. CRC Press, Inc.
- Roche, A., Malandain, G., Ayache, N., 1999. Unifying maximum likelihood approaches in medical image registration. Technical Report, INRIA.
- Roche, A., Pennec, X., Malandain, G., Ayache, N., 2001. Rigid registration of 3D ultrasound with MR images: a new approach combining intensity and gradient information. *IEEE Transactions on Medical Imaging* 20, 1038–1049.
- Rohlfing, T., Maurer, J.C.R., 2004. Modeling liver motion and deformation during the respiratory cycle using intensity-based nonrigid registration of gated MR images. *Medical Physics* 31, 427–432.
- Schneider, U., Pedroni, E., Lomax, A., 1996. The calibration of CT hounsfield units for radiotherapy treatment planning. *Physics in Medicine and Biology* 41, 111–124.
- Sheafor, D.H., Paulson, E.K., Kliewer, M.A., DeLong, D.M., Nelson, R.C., 2000. Comparison of sonographic and ct guidance techniques: does ct fluoroscopy decrease procedure time? *American Journal of Roentgenology* 174, 939–942.
- Stippel, D., Böhm, S., Beckurts, K., Brochhagen, H., Hölscher, A., 2002. Experimental evaluation of accuracy of radiofrequency ablation using conventional ultrasound or a third-dimension navigation tool. *Langenbecks Archives of Surgery* 387, 303–308.
- vanSonnenberg, E., McMullen, W., Solbiati, L. (Eds.), 2005. *Tumor Ablation: Principles and Practice*. Springer.
- Walker, M., Shao, L., Volz, R., 1991. Estimating 3-D location parameters using dual number quaternions. *CVGIP: Image Understanding*, 358–367.
- Wein, W., Khamene, A., 2008. Image-based method for in-vivo freehand ultrasound calibration. In: *SPIE Medical Imaging 2008*, San Diego.
- Wein, W., Pache, F., Röper, B., Navab, N., 2006. Backward-warping ultrasound reconstruction for improving diagnostic value and registration. In: *MICCAI 2006 Proceedings, Lecture Notes in Computer Science*, Springer, pp. 750–757.
- Wein, W., Khamene, A., Clevert, D., Kutter, O., Navab, N., 2007a. Simulation and fully automatic multimodal registration of medical ultrasound. In: *MICCAI 2007 Proceedings*, vol. 4791 of *Lecture Notes in Computer Science*, Springer, pp. 136–143.
- Wein, W., Röper, B., Navab, N., 2007b. Integrating diagnostic B-mode ultrasonography into CT-based radiation treatment planning. *IEEE Transactions on Medical Imaging* 26, 866–879.
- Wood, T., Rose, D., Chung, M., Allegra, D., Foshag, L., Bilchik, A., 2000. Radiofrequency ablation of 231 unresectable hepatic tumors: indications, limitations, and complications. *Annals of Surgical Oncology* 7, 593–600.
- Zagzebski, J.A., 1996. *Essentials Of Ultrasound Physics*. Mosby.
- Zhu, Y., Magee, D., Ratnalingam, R., Kessel, D., 2007. A training system for ultrasound-guided needle insertion procedures. In: *MICCAI 2007 Proceedings*, vol. 4791 of *Lecture Notes in Computer Science*, Springer, pp. 566–574.