

Simulation and Fully Automatic Multimodal Registration of Medical Ultrasound

Wolfgang Wein^{1,2}, Ali Khamene¹, Dirk-André Clevert³, Oliver Kutter²,
and Nassir Navab²

¹ Imaging & Visualization Department
Siemens Corporate Research, Princeton, NJ, USA
{wolfgang.wein,ali.khamene}@siemens.com
² Chair for Computer Aided Medical Procedures (CAMP)
Technische Universität München, Germany
{wein,kutter,navab}@cs.tum.edu
³ Department of Clinical Radiology,
University Hospitals Munich-Grosshadern, Germany
dirk.clevert@med.uni-muenchen.de

Abstract. The fusion of 3D freehand ultrasound with CT and CTA has benefits for a variety of clinical applications, however a lot of manual work 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. The second novelty is a robust similarity measure that assesses the correlation of a combination of multiple signals extracted from CT with ultrasound, without knowing the influence of each signal. This serves as the foundation of a fully automatic registration, which aligns a freehand ultrasound sweep with the corresponding 3D modality using a rigid or an affine transformation model, without any manual interaction. We also present the used initialization, global and local parameter optimization schemes, and validation on abdominal CTA and ultrasound imaging of 10 patients.

1 Introduction

Conventional 2D ultrasound systems can be equipped with position sensing to perform 3D acquisitions of arbitrary size, and to obtain spatial information during the exam. The fusion of such 3D freehand ultrasound imaging with tomographic modalities can, among many other applications, improve the diagnostic value (e.g. for assessment of indeterminate lesions, therefore the term *Diagnostic Fusion*), and integrate anatomic and planning information for interventional navigation of needle procedures. This requires that the target anatomy is precisely registered in ultrasound and the pre-operative modality. Doing so in an automated manner is very challenging, and an active area of research. In [1], 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. Automatic registration on a single kidney CT/US data using Correlation Ratio

as well, here by enhancing the CT intensities with major boundaries, is done in [2]. In [3], 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 [4], 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 a stable registration, yet is computationally efficient at the same time. This has the side effect that the simulation can be used by physicians or sonographers in training to get a feeling for the accessibility and optimal orientations even before the ultrasound exam, or the ultrasound-guided intervention. Besides, a novel similarity measure is developed, which is invariant to missing simulation details, hence having smooth properties and a global maximum at the correct alignment.

2 Simulation of Ultrasound from CT

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 $c = 1540m/s$ in human soft tissue, while a significantly different speed of sound occurs e.g. in air and bone. 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, reflected straight back to the ultrasound transducer depends on the angle:

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

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

The exponent n describes the heterogeneity on the interface, causing the amount of reflection to be more or less narrow around the perpendicular of the tissue interface. We lack detailed physical knowledge from CT, hence we use $n = 1$, as it produces good results and simplifies the equations. The transmitted intensity $t(Z_1, Z_2)$ does not depend on the angle of incidence, if refraction is neglected.

The X-Ray attenuation μ measured by a CT scanner is approximately proportional to the tissue density, see e.g. [5] for a reference table. As tissue density is in turn proportional to acoustic impedance, we can directly derive the an incremental acoustic intensity reflection from it:

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

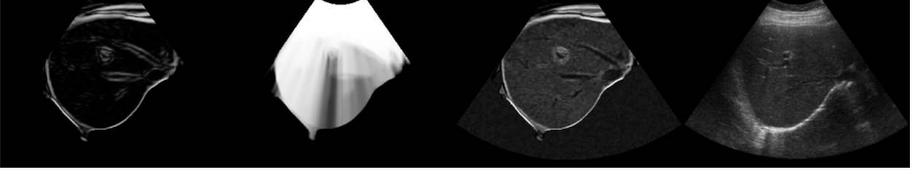


Fig. 1. Simulation of ultrasonic effects from CT, from left to right: Reflection r , transmission t , simulation $r + p$, original ultrasound. 3-dimensional Perlin noise has been added to the occluded part of the simulation.

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

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

$\mu(\mathbf{x})$ is the CT attenuation value at position \mathbf{x} , $\nabla \mu(\mathbf{x})$ its spatial derivative. \mathbf{d} is a unit vector denoting the direction of the ultrasound wave propagation, the scalar multiplication with the normed CT gradient vector yields the angular dependency equivalent to $\cos(\theta)$. The ultrasound wave intensity is reduced according to $t(\mathbf{x})$ at each tissue interface, while $\Delta r(\mathbf{x}, \mathbf{d})$ contributes to the wave intensity detected by the probe. Integrating over this reflection and transmission behavior yields for any depth along a scanline:

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

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 with one parameter a , which amplifies smaller reflections (resembling the Dynamic Range knob on the ultrasound machine), yielding the resulting value of the simulation:

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

For a linear array probe, the integral in equation 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 above row, 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

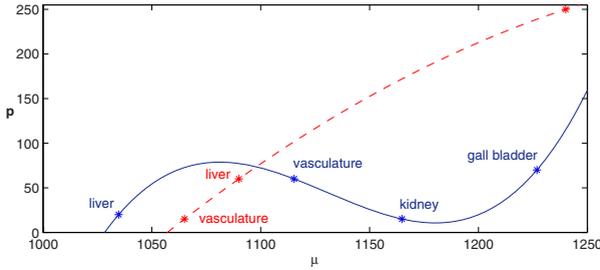


Fig. 2. Intensity mapping p for CT (red, dashed) and portal-venous CTA (blue) soft tissue. Note that the liver-vasculature relation is inverted in the two modalities.

by themselves, based on the microscopic tissue inhomogeneities. There is no simple relationship between tissue echogeneity and CT hounsfield units, therefore we add an intensity mapping $p(\mathbf{x})$ (further described in section 3.2) on a narrow soft-tissue range to the simulated large-scale reflection $r(\mathbf{x})$. Figure 1 depicts the simulation result for a transversal liver image.

3 CT-Ultrasound Registration

3.1 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 vasculature, 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, $n=3$ was used, resulting in 15-22 frames per sweep for registration.

3.2 Idealized Intensity Prior

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 US, they typically produced only a small local optimum. Known approaches for restricting the possible intensity distributions are distance metrics to Joint Histograms learnt from correct registrations (e.g. Kullback-Leibler-Distance), as well as bootstrapping parameters for a polynomial intensity

mapping in the actual registration process itself [1]. 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 define a mapping function $p(\mu)$ based on a number of correspondences (liver tissue, liver vasculature, kidney, gall bladder) between CT/CTA intensities and tissue echogeneity in ultrasound, see figure 2.

3.3 Similarity Measure

In a Correlation Ratio framework, the registration transformation parameters are modified in order to maximize

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

with f denoting the mapping function which estimates the intensities of the image U from the transformed image μ . If a linear mapping $f(\mu) = \alpha\mu + \beta$ is assumed, equation 8 can be directly related to the common Normalized Cross-Correlation (NCC) similarity metric.

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

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

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

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

Therefore the solution is

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

Direct inversion of the symmetric matrix $M^T M$ results in a closed-form solution for the parameters. They are then inserted in equation 8 to yield a novel registration similarity metric, which we denote *Linear Correlation of Linear Combination* (LC^2). It assesses the correlation of ultrasound intensities u_i and a

Table 1. Registration results on 10 patient data sets in terms of the Fiducial Registration Error (FRE) as root mean square (RMS) values in mm

Patient	no. points	manual	pt-based	rigid	affine	remarks
1	8	13.8	9.0	17.0	11.4	strong compression at top
2	7	16.8	10.0	14.4	8.5	
3	11	10.6	8.9	12.0	11.2	10cm renal tumor
5	5	10.0	8.4	15.5	8.7	kidney
6	7	8.0	6.2	10.7	9.9	
7	11	9.1	6.5	10.8	9.3	pt-based reg. visually bad
9	15	4.2	3.5	7.6	6.8	rigid and affine reg. excellent
11	8	11.1	5.6	8.2	8.2	
13	5	11.6	10.7	13.4	12.3	
14	13	6.6	5.4	7.8	8.0	

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 ultrasound image (as NCC), but also independent to how much of the two described physical effects contributes to the image intensities. The 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).

We compute equations 11 and 8 for every ultrasound frame in the set, and use the mean of the results as cost function.

3.4 Optimization Strategy

A rough initial estimate of the orientation is obtained from the tracking setup. The large-scale translation is determined by performing a brute-force scan of the translation space. On the configuration which yields the highest similarity measure value (for a number of similar high results, the one closest to a reasonable preset translation is used), a local optimization of the translation is executed using a Simplex-based non-linear optimizer [6]. Successively, 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. These are the two scaling parameters and the one shearing of the sagittal plane, since respiratory motion mainly causes deformation in that plane [7].

4 Results

An abdominal diagnostic fusion study was performed on 10 patients with various pathology. Our freehand ultrasound system uses an Ascension MicroBird magnetic tracking system with a Siemens Sequoia ultrasound machine and progressive RGS video fed into a PC with frame grabber. The position sensor was

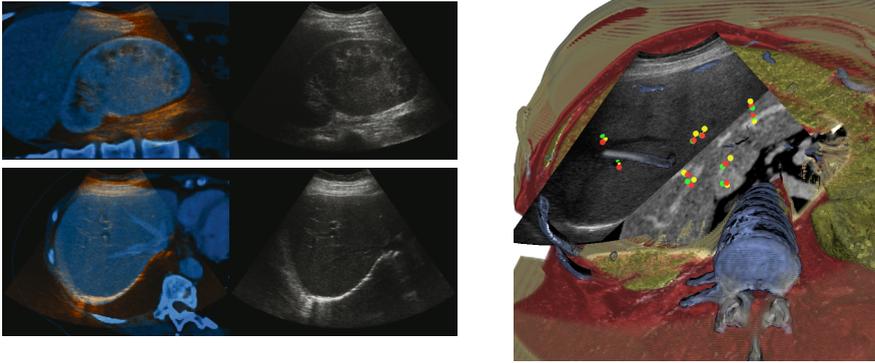


Fig. 3. Left: kidney of patient 5, rigid reg. and liver of patient 6, affine reg. Right: Registered liver of patient 9 with fiducial points (yellow=CT, green=initial US, red=registered US), an oblique CT plane and contextual cutaway volume rendering [9] of CT.

affixed to the transducer using hot-melt adhesive, a method based on [8] was used to determine the calibration. We used the portal-venous phase CTA scans on a dual-source Siemens Somatom Definition scanner, and transversal liver ultrasound sweeps (except for one patient, where a kidney sweep was chosen) for evaluation of the algorithm.

After manually aligning each of the data sets, a physician selected 5-15 point correspondences on anatomical landmarks, including portal & hepatic vein, biliary duct, aorta vena cava and heart atrium. Table 1 lists the RMS distances after manual alignment, point-based rigid registration according to [10], and rigid & semi-affine registration using our methods. The automatic registration converges correctly for all patients with an execution time of ~ 20 seconds. At the initial estimate (before the translation search), the FRE was between $11 - 62\text{mm}$. The errors after automatic alignment are in the same range of the manual ones, but larger than the residual errors after point-based registration. Since all of the registrations seem visually correct (some results are depicted in figure 3), we assume to have a fairly large uncertainty in the definition of point correspondences, especially in cranio-caudal direction. This confirms that manual CT-ultrasound registration is error-prone, as it usually reduces the problem to definition of points on 2D-planes, or manually aligning a single 2D plane (as in use in existing products for interventional CT-US navigation) - not guaranteeing a correct matching in 3D. If affine registration is used, displacements mostly on the top of the images are further reduced (often a large shift of the gall bladder was decreased), accounting for the majority of errors caused by probe pressure, breathing and different patient setup. We expect that the FRE values (all $< 2\text{cm}$) represent an upper bound for a target registration error (TRE) on liver lesions (which we did not define in the scope of this study due to difficult locatability of relevant clinical targets in most of the data).

Regarding the diagnostic value of the study, reading of the registered CT/US data could exclude a number of suspicions on a total of five patients, including partial portal vein thrombosis, acute inflammation of the gall bladder and infiltration of renal cancer into liver tissue.

5 Discussion and Conclusion

We have presented a system for fully automatic alignment of a single freehand ultrasound sweep with CT and CTA data. We expect this to greatly increase the acceptance of multimodal fusion for a number of clinical applications, since it provides a simple workflow and enables more precise registration. Further clinical studies on diagnostic and interventional fusion using the described method are underway. A local variant of the developed LC^2 measure is possible by averaging over smaller overlapping patches, which can further increase the robustness with respect to ultrasound imaging artifacts not covered by the simulation, as well as user adjustments on the ultrasound machine. Besides, LC^2 can easily be extended to handle a larger number of signals from both modalities. Real-time compensation of respiratory motion and deformable mapping techniques will be investigated as well, based on the proposed methods.

References

1. Roche, A., Pennec, X., Malandain, G., Ayache, N.: Rigid registration of 3D ultrasound with MR images: a new approach combining intensity and gradient information. *IEEE Trans. Med. Imag.* 20, 1038–1049 (2001)
2. Leroy, A., Mozer, P., Payan, Y., Troccaz, J.: Rigid registration of freehand 3D ultrasound and CT-Scan kidney images. In: Barillot, C., Haynor, D.R., Hellier, P. (eds.) *MICCAI 2004*. LNCS, vol. 3216, p. 837. Springer, Heidelberg (2004)
3. Penney, G., Blackall, J., Hamady, M., Sabharwal, T., Adam, A., Hawkes, D.: Registration of freehand 3D ultrasound and magnetic resonance liver images. *Medical Image Analysis* 8, 81–91 (2004)
4. Wein, W., Röper, B., Navab, N.: Automatic registration and fusion of ultrasound with CT for radiotherapy. In: Duncan, J.S., Gerig, G. (eds.) *MICCAI 2005*. LNCS, vol. 3750, pp. 303–311. Springer, Heidelberg (2005)
5. Schneider, U., Pedroni, E., Lomax, A.: The calibration of CT hounsfield units for radiotherapy treatment planning. *Phys. Med. Biol.* 41, 111–124 (1996)
6. Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: *Numerical Recipes in C*, 2nd edn. CRC Press, Inc. (1992)
7. Rohlfing, J.T., Maurer, C.R.: Modeling liver motion and deformation during the respiratory cycle using intensity-based nonrigid registration of gated MR images. *Medical Physics* 31, 427–432 (2004)
8. Rousseau, F., Hellier, P., Barillot, C.: Confhustus: A robust and fully automatic calibration method for 3D freehand ultrasound. *Medical Image Analysis* 9 (2005)
9. Burns, M., Haidacher, M., Wein, W., Viola, I., Groeller, E.: Feature emphasis and contextual cutaways for multimodal medical visualization. In: *EuroVis 2007 Proceedings* (2007)
10. Walker, M., Shao, L., Volz, R.: Estimating 3-D location parameters using dual number quaternions. *CVGIP: Image Understanding*, 358–367 (1991)