

AUTOMATIC NON-LINEAR MAPPING OF PRE-PROCEDURE CT VOLUMES TO 3D ULTRASOUND

Wolfgang Wein¹, Oliver Kutter¹, Andre Aichert², Darko Zikic², Ali Kamen¹, Nassir Navab²

¹Imaging & Visualization Department
Siemens Corporate Research, Princeton, NJ, USA

²Computer Aided Medical Procedures (CAMP)
Technische Universität München, Germany

ABSTRACT

Multi-modality alignment of CT and ultrasound adds value to diagnostic examinations, as well as treatment planning and execution of various clinical procedures. Particularly automatic image-based alignment of such data is challenging, mostly because both modalities have very different imaging physics and characteristics. We present a method for dense-field deformable registration of CT and 3D ultrasound. Compared to global (rigid) alignment, this is more difficult to solve, because modality-specific difference in local anatomic appearance can result in incorrect displacements. We use a simulation of ultrasonic effects based on CT information, taking the current estimate of the deformation field into account to properly address orientation-dependent imaging artifacts. This is combined with a robust multi-channel local similarity metric, driving a variational registration framework. Because of the high computational demand, an efficient GPU-based implementation is used. Preliminary results are shown on data from a number of hepatic cancer patients. To our knowledge, this is the first time that a non-linear mapping of CT and 3D B-mode ultrasound is established in a computationally practical and fully automatic manner.

Index Terms— Deformable Registration, CT, Ultrasound

1. INTRODUCTION

Image registration seeks to spatially align different images to increase the information content, or expand the field of view. In the medical domain, multi-modality registration can provide important clues, as different exams on the same patient often provide complementary information. Furthermore, aligning a pre-operative data set with intra-operative imaging of another modality enables advanced image-based guidance and navigation systems.

In this work, we address the problem of automatically computing a deformable mapping of a CT scan to 3D ultrasound of the same patient, where the focus is on liver anatomy.

To our knowledge, there is no prior art which addresses

direct image-based registration of CT and B-mode ultrasound with a non-linear deformation model. Somewhat related, in [1], sparse Doppler ultrasound data for liver surgery is registered to CT, using a combination of point correspondences and an image-based term based on normalized gradient fields. In [2], MRI is registered to 3D freehand ultrasound, without reconstructing the latter into a Cartesian volume, using a B-spline based approach.

2. METHOD

We seek to find a displacement function $u : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, which defines a deformation $\varphi = Id + u$ which warps the template CT image T , such that it is aligned to the 3D ultrasound reference image R . Thus, the problem becomes a minimization of the following energy

$$\hat{u} = \arg \min_u \mathcal{D}(R, T \circ (Id + u)) \quad (1)$$

in which \mathcal{D} is an appropriate difference measure. One important property of the displacement u is its regularity, which can be achieved in different ways. One possibility is to include a regularization term directly into the energy formulation, where a wide range of different terms and optimization approaches is possible, compare for example [3]. Further regularization can be achieved by parameterizing the deformation, e.g. by B-Splines, compare [4]. In this work, we take a different approach which does not require an explicit regularization term in the energy formulation, compare [5]. The deformation is constrained approximately to the manifold of diffeomorphisms by computing the result as a consecutive composition of small update fields. The updates are computed by calculating the gradient of the energy function and smoothing it by applying a low-pass filter. This approach has the advantage that the optimization can be implemented very efficiently if fast filtering methods are available.

For the difficult case of CT-Ultrasound registration, we will now incorporate a number of additions that are not part of the standard formulation.

2.1. Ultrasound Simulation from CT

To enable a better assessment of structural alignment in CT and ultrasound, we simulate ultrasonic imaging effects from CT, as previously described in [6]. Let $\mu : \Omega_T \mapsto \mathbb{R}$ be the volume of original CT intensity values (i.e. Hounsfield units). Under the assumption that the acoustic impedance of human tissue is proportional to its density, which in turn relates to the CT Hounsfield measurements, ultrasonic reflections at tissue interfaces can be estimated:

$$T_1(\mathbf{x}) = \log \left[aI_0 \exp \left(- \int_0^{\lambda_x} \left(\frac{|\nabla \mu(\mathbf{x}_0 + \lambda \mathbf{d})|}{2\mu(\mathbf{x}_0 + \lambda \mathbf{d})} \right)^2 d\lambda \right) \right. \\ \left. (\mathbf{d}^T \nabla \mu(\mathbf{x})) \frac{|\nabla \mu(\mathbf{x})|}{(2\mu(\mathbf{x}))^2} + 1 \right] \frac{1}{\log(a+1)} \quad (2)$$

In words, for a voxel location \mathbf{x} , the reflection T_1 can be computed by integrating over reflection and transmission along an ultrasonic scanline originating at \mathbf{x}_0 with direction \mathbf{d} . It is embedded in a log-compression with parameter a . This creates a reasonable simulation of ultrasound reflections, including shadowing effects (i.e. after a strong reflection the remainder of that scanline will have no reflection). The second imaging effect is tissue echogeneity, which results from the scattering behavior of tissue inhomogeneities, that are smaller than the ultrasound pulse wavelength. It is approximated by a heuristic mapping of the CT soft tissue intensity range, and modulated with the ultrasonic transmission from the integral in equation 2, in order to clear its value if shadowing has occurred. We define this second term as $T_2(\mathbf{x})$. Because those simulated values T_1 and T_2 are derived from line integrals through the CT volume, they implicitly depend on the entire deformation field φ . For T_1 , this results to

$$T_1(\varphi, \mathbf{x}) = \log \left[aI_0 \exp \left(- \int_0^{\lambda_x} \left(\frac{|\nabla \mu(\varphi(\mathbf{x}_0 + \lambda \mathbf{d}))|}{2\mu(\varphi(\mathbf{x}_0 + \lambda \mathbf{d}))} \right)^2 d\lambda \right) \right. \\ \left. (\mathbf{d}^T \nabla \mu(\varphi(\mathbf{x}))) \frac{|\nabla \mu(\varphi(\mathbf{x}))|}{(2\mu(\varphi(\mathbf{x})))^2} + 1 \right] \frac{1}{\log(a+1)}$$

Back to the notation introduced before, our template image becomes a two-channel representation $T : \Omega_T \mapsto \mathbb{R}^2$, whose voxel values are implicitly dependent on the whole deformation field, i.e.

$$T(\mu, \varphi, \mathbf{x}) \doteq \Theta_{\mu, \varphi}(\mathbf{x}) \quad (3)$$

Θ is therefore our "ultrasound simulation operator" which computes T_1 and T_2 given a CT volume μ and deformation field φ .

2.2. Similarity Metric

We use a local dissimilarity metric that is evaluated in the neighborhood of every grid location \mathbf{x} . Typical examples of such measures are local normalized cross-correlation (LNCC) or local Mutual Information. See [7] for the mathematical foundation in the context of variational registration. One common aspect of those (dis-)similarity measures is that they

assume a certain relationship between the image intensities, which can often be modeled as a functional dependency. In the case of cross-correlation it is a linear mapping between the image intensities, in the case of generalized correlation ratio it might e.g. be a polynomial mapping [8]. This intensity mapping is usually expressed with respect to a number of unknown parameters, which can be intrinsically solved for during computation of the similarity (as in cross-correlation), or during a separate optimization. In Computer Vision, this mapping is usually denoted as *photometric transformation*.

In [6], the LC^2 similarity metric is proposed, which is independent of how much the simulated channels T_1 and T_2 locally contribute to the actual ultrasound intensity R . However, the parameters describing this intensity contribution are solved in a least-squares sense for every local evaluation of the metric. This makes it difficult to analytically derive its gradient; thus we use the following approximation for \mathcal{D} , which proves to carry the sufficient information to successfully drive the registration process:

$$\mathcal{D}(R, T, \varphi) = -\mathcal{D}_{LCC}(R, T_1 \circ \varphi) - \mathcal{D}_{LCC}(R, T_2 \circ \varphi) \quad (4)$$

where \mathcal{D}_{LCC} is the local cross-correlation (LCC) measure, defined in [7] for images I_1 and I_2 as

$$\mathcal{D}_{LCC}(I_1, I_2, \varphi, \mathbf{x}) = \frac{v_{1,2}(\varphi, \mathbf{x})^2}{v_1(\mathbf{x})v_2(\varphi, \mathbf{x})} \quad (5)$$

where $v_1(\mathbf{x})$ is the variance of I_1 , $v_2(\varphi, \mathbf{x})$ is the variance of $I_2 \circ \varphi$, and $v_{1,2}(\varphi, \mathbf{x})$ is the covariance of I_1 and $I_2 \circ \varphi$. The major idea of the LCC measure is that the joint probability of the two images I_1 and I_2 , from which the means and variances are computed, is weighted locally by a Gaussian kernel G_γ with standard deviation γ . Therefore the variances and the covariance are dependent on the location in space \mathbf{x} .

The rationale behind replacing the linear combination from [6] with equation 4, is that the change of the individual LCC metrics depends on how much content each simulated channel T_1 , T_2 provides. For example, if in a particular voxel neighborhood, no reflection term T_1 contributes to the similarity, its influence on the displacement update will be minor. Therefore, the sum of two individual LCC measurements can account for the unknown sum of ultrasound reflection and echogeneity. From Eq. (4), the gradient of \mathcal{D} is

$$\nabla \mathcal{D}(R, T, \varphi) = -\nabla \mathcal{D}_{LCC}(R, T_1 \circ \varphi) - \nabla \mathcal{D}_{LCC}(R, T_2 \circ \varphi)$$

The gradient of the local cross-correlation with respect to φ , at position in space \mathbf{x} , is given in the following form:

$$\nabla \mathcal{D}_{LCC}(\varphi)(\mathbf{x}) = L(I_1, I_2, \varphi, \mathbf{x}) \nabla I_2(\varphi(\mathbf{x})) \quad (6)$$

For the case of LCC, the term L is defined as

$$L(I_1, I_2, \varphi, \mathbf{x}) = G_\gamma \star \frac{-2}{\mathcal{G}_\gamma} \left[\frac{v_{1,2}(\varphi, \mathbf{x})}{v_2(\varphi, \mathbf{x})} \left(\frac{I_1(\mathbf{x}) - \overline{I_1(\mathbf{x})}}{v_1(\mathbf{x})} \right) \right. \\ \left. - \mathcal{D}_{LCC}(I_1, I_2, \varphi, \mathbf{x}) \left(\frac{I_2(\varphi(\mathbf{x})) - \overline{I_2(\varphi, \mathbf{x})}}{v_2(\varphi, \mathbf{x})} \right) \right]$$

where \star is the convolution operator, \mathcal{G}_γ is the integral of the local Gaussian kernel inside the image domain, $\overline{I_1(\mathbf{x})}$ is the local mean intensity of I_1 at \mathbf{x} , and $\overline{I_2(\varphi, \mathbf{x})}$ is the local mean intensity of $I_2 \circ \varphi$ at \mathbf{x} .

By substituting I_1 with the reference R and I_2 with T_1 and T_2 , respectively, the gradient of \mathcal{D} can be computed.

2.3. Optimization

The solution to the optimization problem is found by iteratively computing an update du to the initial displacement estimate $u^{(i+1)}$ over iterations i in the following form

$$\begin{aligned} du &\leftarrow G_\sigma(\nabla\mathcal{D}(R, T \circ (\text{Id} + u^{(i)}))) & (7) \\ u^{(i+1)} &\leftarrow u^{(i)} \circ (\text{Id} + \tau du) & (8) \end{aligned}$$

Here, G is a Gaussian filtering operation with standard deviation σ , $\nabla\mathcal{D}(R, T \circ (\text{Id} + u^{(i)}))$ is the derivative of the similarity measure, id is the identity transformation, and the positive scalar τ is the update magnitude control parameter.

2.4. Implementation

The approach described above is computationally expensive, because it requires a complete simulation of 3D ultrasound from CT in every iteration. In order to exploit the potential for concurrent programming, as well as hardware support for all volume resampling operations, we implemented this method on graphics hardware. The OpenGL graphics library and shading language (GLSL) are used [9]. In every major iteration of the algorithm, a 3D freehand ultrasound sweep is simulated, followed by a reconstruction into a Cartesian volume [10]. Then, the displacement field update is computed per voxel, smoothed by a low-pass filter, and the deform image is recomputed by warping the source image. Simulation, reconstruction and registration are computed on the GPU, which completely avoids costly transfers in order to communicate results between the algorithms. Convolution performance of current graphics hardware is texture-read bound and can be even worse than recursive convolution on multi-core CPUs for larger kernels. Therefore, the LCC computation utilizes a semi-recursive convolution along the 3D texture slice stack to optimize performance.

3. RESULTS

We evaluated the registration on data from 12 different patients with liver lesions (with different fidelity and patient setup). Each patient data comprised a portal-venous phase contrasted CT scan and a 3D ultrasound volume of the liver. The ultrasound has been recorded using a magnetically tracked 2D probe, swept in a transversal orientation over the patient’s liver during breath-hold.

Our registration converges in 10–30 iterations, depending on the complexity of the recovered displacement. A typical

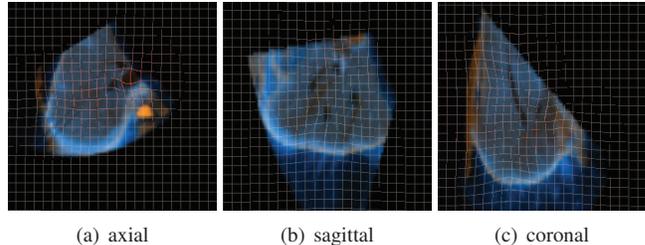


Fig. 1. Color overlay of registration result (US simulated from CT $T_1 + T_2$ in brown, reference US blue) and resulting deformation grid. The patient was in supine position during the ultrasound exam.

computation time per iteration is $\approx 240ms$, which consists to roughly 18% of the ultrasound simulation operator Θ , 49% compounding into a Cartesian grid, and 33% computing \mathcal{D} and the displacement update, the overall 3D voxel size of the system is $\tilde{1}28^3$. The entire execution time is 2–10s, depending on the actual abortion criteria, and the voxel size / slice count of the data sets. After the first few iterations, the size of the displacement updates becomes small enough that a re-simulation of ultrasound from CT is not necessary, as it does not significantly influence the orientation-dependant artifacts. Each iteration then encounters a 3-fold speedup, resulting in a possible overall computation time of less than one second. In order to validate registration accuracy, a physician manually identified anatomical landmarks and clinical targets in both modalities, resulting in 5–16 point correspondences for each patient. A rigid motion computed from those points comprises the “Ground Truth”. Initialized with this, a rigid image-based registration [6] was launched, increasing the RMS error, but better aligning large-scale peripheral structures. Our deformable algorithm was executed from this configuration, always with the same parameters (σ , convergence criteria, simulation parameters etc.) to demonstrate possible automation. The average RMS error at the Ground Truth is 5.5mm, after rigid registration 8.8mm, after deformable registration 8.2mm (see table 1). Our algorithm improved the error from rigid registration in 9/12 cases, on average 60% of the point correspondences were improved. Visual alignment improved in all but one case (LPO setup with strongly misaligned hepatic vasculature after rigid registration).

4. DISCUSSION

We have presented an approach that solves the problem of automatically computing a dense-field deformable mapping of CT and 3D ultrasound of the same patient. This becomes possible by using a carefully chosen combination of algorithms and components. It includes a modality mapping operator Θ , simulating ultrasound from CT - which goes beyond typical photometric transformation approaches that are common in

GT	Rigid	Def.	Points	Setup
3.51	10.3	9.6	12	LPO
4.29	10.9	11.2	6	LPO
8.02	10.3	8.3	9	decubitus
9.25	11.1	11.2	11	supine
4.03	5.43	5.39	11	supine
3.31	4.32	3.78	11	supine
3.68	5.10	5.51	7	supine
3.23	8.97	7.99	7	supine
8.34	12.3	11.2	16	LPO
5.25	6.26	5.54	8	supine
6.15	9.89	9.15	8	supine
6.42	10.3	9.2	6	LPO

Table 1. Root-Mean-Square (RMS) errors of point correspondences at the Ground Truth (GT) alignment, as well as after rigid and deformable registration, for 12 patients. Also listed are the number of points, and the patient setup during the ultrasound exam (LPO = left portal oblique).

the field. A two-component similarity measure based on local cross-correlation is used in a variational formulation. Powerful concurrent GPU-based programming techniques are used to make the approach feasible in terms of computation time.

We showed results on 12 data sets. Our algorithm offers both visually and quantitatively convincing deformable mapping, effectively compensating for differences in patient setup, respiration and other sources of organ motion. Given that applying a full deformation on a pre-operative data set might be undesirable in clinical practice, the presented algorithm can also be used to derive an improved rigid or global-affine registration around the target anatomy (avoiding the problem of multiple local optima, caused by the fact that those transformation models can never truly match the data).

Our method still sometimes produces locally inaccurate deformations, in areas where the simulation does not resemble the real ultrasound images. Future work therefore includes refining the simulation model, as well as automatically detecting areas lacking anatomical information in one of the modalities.

5. REFERENCES

- [1] N. Papenberg, T. Lange, J. Modersitzki, P.M. Schlag, and B. Fischer, "Image registration for CT and intra-operative ultrasound data of the liver," in *SPIE Medical Imaging*, San Diego, California, USA, February 2008.
- [2] R. Brooks, D.L. Collins, X. Morandi, and T. Arbel, "Deformable ultrasound registration without reconstruction," in *MICCAI 2008 Proceedings*, 2008, vol. 5242 of *Lecture Notes in Computer Science*, pp. 1023–1031.
- [3] Jan Modersitzki, *Numerical methods for image registration*, Oxford University Press, 2004.
- [4] D. Rueckert, L.I. Sonoda, C. Hayes, D.L.G. Hill, M.O. Leach, and D.J. Hawkes, "Nonrigid registration using free-form deformations: application to breast mr images," *Medical Imaging, IEEE Transactions on*, vol. 18, no. 8, pp. 712–721, Aug. 1999.

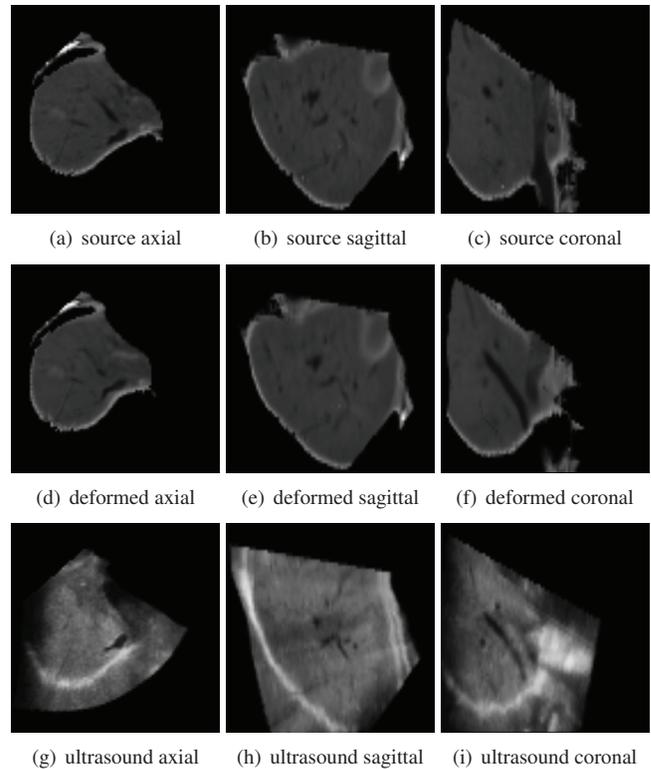


Fig. 2. Cross-sections of original simulated US $T_1 + T_2$, deformed simulated US and reference US for a patient. The ultrasound exam was performed with the patient in left portal oblique (LPO) position.

- [5] C. Chefd'hotel, G. Hermosillo, and O. Faugeras, "Flows of diffeomorphisms for multimodal image registration," in *IEEE International Symposium on Biomedical Imaging*, 2002, pp. 753–756.
- [6] W. Wein, S. Brunke, A. Khamene, M.R. Callstrom, and N. Navab, "Automatic ct-ultrasound registration for diagnostic imaging and image-guided intervention," *Medical Image Analysis*, vol. 12, pp. 577–585, Oct. 2008.
- [7] G. Hermosillo, *Variational Methods for Multimodal Image Matching*, Ph.D. thesis, Universite de Nice - Sophia Antipolis, 2002.
- [8] A. Roche, X. Pennec, G. Malandain, N. Ayache, and S. Ourselin, "Generalized correlation ratio for rigid registration of 3D ultrasound with MR images," Tech. Rep., INRIA, July 2000.
- [9] O. Kutter, R. Shams, and N. Navab, "Visualization and GPU-accelerated simulation of medical ultrasound from CT images," *Computer Methods and Programs in Biomedicine*, vol. 94, no. 3, pp. 250–266, June 2009.
- [10] A. Karamalis, W. Wein, O. Kutter, and N. Navab, "Fast hybrid freehand ultrasound volume reconstruction," in *SPIE Medical Imaging*, Orlando, Florida, USA, February 2009.