

# Multi-modal Registration Based Ultrasound Mosaicing

Oliver Kutter<sup>1</sup>, Wolfgang Wein<sup>2</sup>, and Nassir Navab<sup>1</sup>

<sup>1</sup> Computer Aided Medical Procedures, TUM, Munich, Germany  
{kutter,navab}@cs.tum.edu

<sup>2</sup> Siemens Corporate Research, Princeton, NJ, USA  
wolfgang.wein@siemens.com

**Abstract.** Recent US systems allow the real-time acquisition of volume data, either by freehand 3D techniques or novel transducer hardware. However, the acquisition of large volumes is limited by the field of view of the US transducer and anatomical scanning windows into the patient. Mosaicing of several 3D US scans has been proposed to generate large US volumes. Still, US imaging specific characteristics and artifacts make it challenging to create high quality mosaics. For many clinical cases, especially interventions, additional high quality CT data is available. In this paper we present a novel multi-variate, multi-modal 3D US registration and mosaicing approach, which reduces the effects of ultrasound imaging artifacts on mosaic quality, by incorporating information from co-registered CT.

## 1 Introduction

3D US imaging is becoming more and more frequently used in today's clinical workflows. Recent US systems allow direct real-time acquisition of 3D volume data. Either freehand 3D US, where a position and orientation (pose) sensor is attached to a conventional 2D US probe, or direct 3D acquisition capable US transducers are used. However the acquisition of large volume data is today still limited by the US transducer field of view and restricted scanning windows into the patient. Analog to 2D US image mosaicing, US volume mosaicing has been proposed to generate large US volumes from several acquired 3D US datasets.

Such 3D mosaic volumes of entire organs help to improve diagnostic confidence, e.g. by reducing scanning orientation induced artifacts, and transcend the role of ultrasound in radiology by providing self-contained 3D data for the reading room. For guidance of minimally-invasive interventions, such mosaics provide global anatomical reference, and allow for automatic mono- and multi-modal fusion with pre-operative data. US imaging specific artifacts, e.g. noise, occlusions, speckle, make it challenging to design robust similarity measures. By using the enriched spatial and artifacts reduced information of US mosaics the registration problem is becoming less ill-posed. Besides, comparison of 3D US at different stages before, during and after the procedure, allow to better assess treatment outcome. Additionally, 3D US mosaics could also be used as databases for generating simulated US images for training and education of physicians [1].

## 1.1 Problem Statement

3D US mosaicing is a two step process. First the unknown registration parameters  $[\hat{\mathbf{p}}_0, \dots, \hat{\mathbf{p}}_N]$  of all individual volumes have to be computed. The registration problem is solved by minimization of a cost function

$$[\hat{\mathbf{p}}_0, \dots, \hat{\mathbf{p}}_N] = \arg \min_{[\mathbf{p}_0, \dots, \mathbf{p}_N]} \frac{\sum_{k=0}^N \sum_{l=0}^N \omega_{(k,l)} \cdot \sum_{\mathbf{v} \in \Omega} \Gamma(I_k[T^{-1}(\mathbf{p}_k) \cdot \mathbf{v}], I_l[T^{-1}(\mathbf{p}_l) \cdot \mathbf{v}])}{\sum_{k=0}^{N-1} \sum_{l=k+1}^N \omega_{(k,l)}} \quad (1)$$

defined by the sum of the pair-wise cost functions  $\Gamma$  for all pairs  $(I_k, I_l)$  over all elements  $v$  in the overlap domain  $\Omega$ , where  $T^{-1}(\mathbf{p}_i)$  maps an element in overlap coordinates into the  $i$ -th volume.  $\omega_{(k,l)}$  is a weighting factor for the contribution of the costs for pair  $(I_k, I_l)$  to the global costs. If one chooses  $\omega_{(k,l)} = 1$  for all pairs the total cost is equal to the average of the individual costs. Another option is to choose  $\omega_{(k,l)}$  equal to the pair-wise volume overlap, as proposed by Wachinger et al [2] to introduce a weighting favoring larger overlaps over smaller ones.

For solving the registration problem several approaches exist. They can be classified by the use of either direct or indirect cost functions, and pair-wise and simultaneous optimization approaches. Indirect approaches first pre-process the 3D US data to extract features for registration, in contrast direct approaches directly use the complete image intensity information. Pair-wise approaches decompose the global registration problem into several sequentially executed pair-wise registration problems, which are then used to compute the global registration by propagation of transformations. Simultaneous approaches directly minimize the cost function by simultaneously solving for all parameters and are more stable towards pair-wise local minima. However, they also greatly increase the computational effort per cost function evaluation. For symmetric cost functions, Eq.1 can be simplified to

$$\frac{\sum_{k=0}^{N-1} \sum_{l=k+1}^N \omega_{(k,l)} \cdot \sum_{\mathbf{v} \in \Omega} \Gamma(I_k[T^{-1}(\mathbf{p}_k) \cdot \mathbf{v}], I_l[T^{-1}(\mathbf{p}_l) \cdot \mathbf{v}])}{\sum_{k=0}^{N-1} \sum_{l=k+1}^N \omega_{(k,l)}} \quad (2)$$

which reduces the evaluation costs significantly.

Once the registration problem is solved, the actual per pixel mosaic values have to be computed from the individual US intensities of different volumes. Both, the registration and mosaic steps are challenging to solve for US data due to the characteristics of US imaging itself. The intensity of the same anatomical region varies in different US scans for multiple reasons. Variations in direction of the US sound waves, give rise to occlusion and shadowing artifacts. Hypoechoic regions can cause negative shadows in scan direction. Motion, either induced by pressure of the transducer onto the patient, internal organ motion, and of course breathing motion cause geometric variations in the images. Furthermore, US system internal signal processing, noise, speckle and operator dependent imaging control settings cause variations between different scans.

Several recent works have dealt with the problem of 3D US mosaicing. Gee et al [3] used a direct pair-wise registration approach. They reduce the complexity of the 3D/3D registration problem to a 2D/2D registration at volume intersection planes. Poon et al [4] proposed two block-based methods for improving the local alignment in the overlap volumes to compensate for the errors introduced by position trackers. Wachinger et al. [2] investigated registration strategies, pair-wise vs. simultaneous, for US mosaicing on a single phantom dataset. Simultaneous approaches proved to be superior to pair-wise ones, however resulting in high computational costs. Ni et al [5] chose an indirect pair-wise registration approach, using SIFT feature extraction. They evaluated their approach on an abdominal phantom and 3D US scans of human liver.

In this work we introduce a new simultaneous, direct, and multi-modal 3D US mosaicing approach. We chose not to use an indirect approach as reliable extraction of features from US data is inherently ill-posed in general for scans with varying field of view, scan direction, and unknown volume overlap. For many clinical scenarios a high quality CT scan of the patient is available. Our method uses the additional information from CT to improve the US/US registration by simultaneous registration of US to US and US to CT data. US/US registration problems caused by small overlap, lack of registration relevant features in the overlap volume, can be compensated by the additional information from CT. We use a variation of the method for CT/US fusion introduced by Wein et al [6] to estimate the registration parameters by simultaneously computing the registration of all US volumes to each other and to the CT scan. A simultaneous optimization was chosen instead of a pair-wise one to avoid accumulation of errors and improve overall stability of the registration. To reduce the computational costs we implemented the complete registration algorithm on GPU stream processors, greatly accelerating the multi-variate cost function evaluation compared to a CPU-based implementation.

## 2 Methods

**Registration Framework.** We use a GPU accelerated registration framework, supporting pair-wise, and simultaneous intensity based registration. Several, mono and multi-modal similarity measures (c.f. Table 1) have been implemented using OpenGL/GLSL<sup>1</sup>. We use global affine transformations, as rigid transformations are not sufficient to cover scale variations and shearing due to motion, scan conversion, compounding, shape of US field of view in the individual scans. A Downhill Simplex optimizer was chosen to optimize the similarity measure and determine the registration parameters. We added a transformation penalty term to the similarity measure to penalize physical impossible or unlikely transformations for US scans with a known pose guess from tracking. For US data with unknown pose, either a rough manual pre-positioning and relaxed penalty condition is used, or no transformation restriction at all.

<sup>1</sup> OpenGL - Open Graphics library <http://www.opengl.org>  
GLSL - OpenGL's built-in high level shading language.

**Table 1.** Used similarity measures. Left colun, pair-wise, right column simultaneous version. (Row 1) n = 1 SAD - Sum of absolute differences, n = 2, SSD - Sum of squared differences (Row 2) NCC - Normalized Cross Correlation, (Row 3) LNCC - Local NCC in a  $k^3$  neighbourhood.

$$\begin{aligned} \frac{1}{N} \sum_{k=1}^N |I_k - I_l|^n & \qquad \frac{\sum_{k=0}^{N-1} \sum_{l=k+1}^N \sum_{\mathbf{v} \in \Omega} |I_k - I_l|^n}{\sum_{k=0}^{N-1} \sum_{l=k+1}^N \omega(k,l)} \\ NCC(I, J) &= \frac{\sum_{k=1}^N (I_k - \bar{I})(J_k - \bar{J})}{N \sigma_I \sigma_J} \qquad \frac{\sum_{k=0}^{N-1} \sum_{l=k+1}^N \omega(k,l) \sum_{\mathbf{v} \in \Omega} NCC(I_k, I_l)}{\sum_{k=0}^{N-1} \sum_{l=k+1}^N \omega(k,l)} \\ LNCC(I, J, k) &= \frac{1}{N} \sum_{nb=k}^N NCC(I, J)^k \qquad \frac{\sum_{k=0}^{N-1} \sum_{l=k+1}^N \sum_{\mathbf{v} \in \Omega} LNCC(I_k, I_l)}{\sum_{k=0}^{N-1} \sum_{l=k+1}^N 1} \end{aligned}$$

**2.1 Multi-modal Mosaicing**

The above described registration framework directly allows for mono-modal mosaicing of 3D US data. We incorporate the CT information into the registration algorithm by a variation of the method of Wein et al [6]. In every iteration the major US reflections are simulated from CT [7] and reconstructed into a cartesian volume [8] on the GPU. As similarity measure  $\varphi$  we chose the weighted sum of the multi-variate normalized cross correlation of all 3D US data, and real 3D US and simulated reflections (cf. Eq. 3).

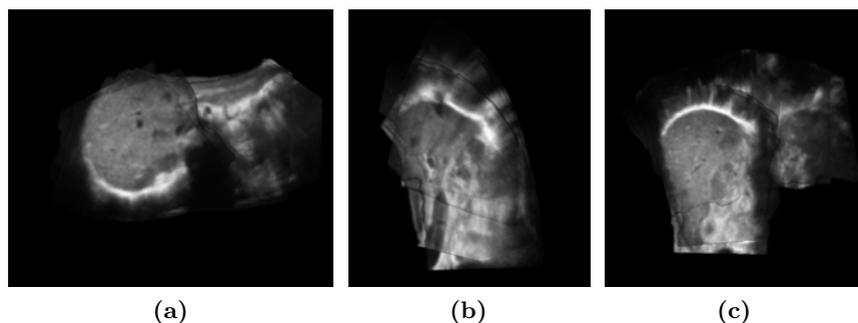
$$\begin{aligned} \varphi &= \frac{\sum_{k=0}^{N-1} \sum_{l=k+1}^N \omega(k,l) \cdot \sum_{\mathbf{v} \in \Omega} \mathbf{NCC}(US_k, US_l)}{\sum_{k=0}^{N-1} \sum_{l=k+1}^N \omega(k,l)} \\ +\alpha & \frac{\sum_{k=0}^{N-1} \sum_{l=k+1}^N \omega(k,l) \cdot \sum_{\mathbf{v} \in \Omega} \mathbf{NCC}(\theta(CT, US_k), US_l)}{\sum_{k=0}^{N-1} \sum_{l=k+1}^N \omega(k,l)} \end{aligned} \tag{3}$$

where  $\theta(CT, US_k)$  is the US simulation operator, simulating the  $k$ -th US volume from CT. For final refinements after registration we exchange the global NCC in  $\varphi$  with a Local NCC. Local NCC is not selected right from the start as NCC is stable enough for aligning the major features in US/US and CT/US data from the initial pose.

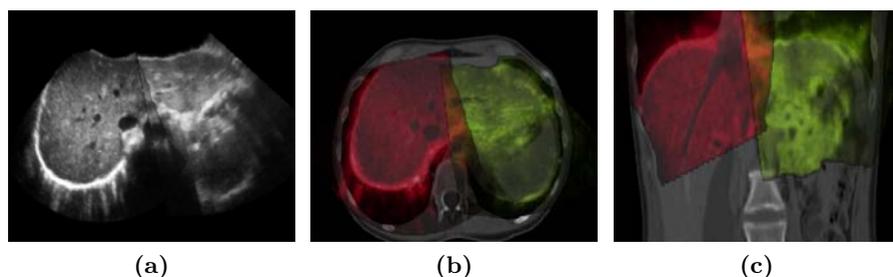
We chose this measure as the LC2 measure proposed by Wein et al [6] for CT/US fusion proved numerically instable on recent GPUs using GLSL due to lack of double precision floating point support. However, the simulated major reflections provide sufficient information for stable CT/US registration for US scans with large tissue interfaces.

**3 Experiments and Results**

We evaluated our 3D US mosaicing framework for phantom and real clinical 3D freehand US datasets of human liver from multiple patients. Both mono- and multi-modal mosaicing were tested. The stability of the simultaneous US/US registration was evaluated on four 3D US scans of a baby phantom. The US probe was moved in cranio-caudal direction over the phantom, without a major view



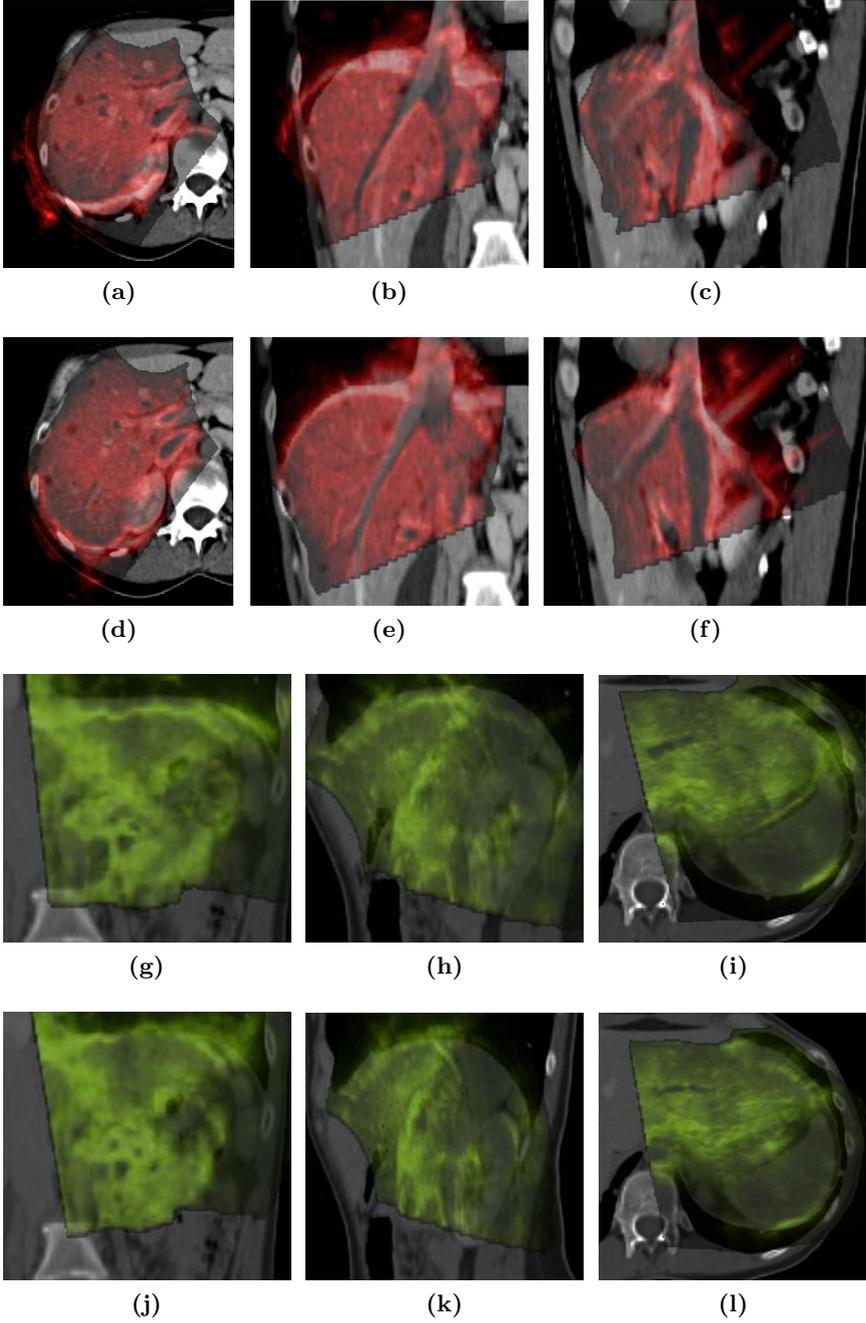
**Fig. 1.** Axial along mosaic volume generated 2D MPRs from five freehand 3D US scans of human liver



**Fig. 2.** (a) Mosaic from two volumes with very small overlap in the middle. Mosaicing works only if CT information (b,c) is used to constrain the registration. Use of mono-modal approach will lead to breakup of the volume configuration.

direction change. Therefore the intensity variation in the scans is mostly due to noise. An optimization of translation and rotation parameters was performed 100 times with random offsets of  $\pm 10$ mm and  $\pm 5$  degrees from a defined ground truth pose. SSD was used as similarity measure. Table 2a depicts the standard deviations of the translation and rotational parameters for the phantom and patient dataset.

The same stability study was performed for multiple real clinical datasets. Fig. 1 shows an example of an US mosaic created from four 3D freehand US scans of the liver region. For the clinical data NCC was selected as similarity measure. Table 3a depicts the results of the stability study. The results show that direct simultaneous registration performs as well on phantom as on real patient data. Tables 2b and 2b depict the performance of the framework for both mono and multi-modal registration. Table 2b shows an impressive speedup of more than 400 times of the GPU version versus a single-threaded CPU version for the same data and initial registration parameter settings. The used computer was equipped with an Intel Core Duo processor (2.6Ghz), 4Gbyte RAM and a NVIDIA Geforce GTX 280 with 1GByte of VRAM.



**Fig. 3.** Comparison of CT/US alignment before (a-c,g-i) and after registration (d-f,j-l) for the two US volumes with small overlap depicted in Fig. 2

**Table 2.** Simultaneous mono-modal registration experiment results for four US phantom scans. (a) Standard deviations of translation and rotation parameters for random study from defined ground truth pose. (b) GPU vs. CPU performance comparison for different volumes resolutions.

(a)							(b)		
Vol.	tx [mm]	ty [mm]	tz [mm]	$\alpha$ [rad]	$\beta$ [rad]	$\gamma$ [rad]		GPU [sec]	CPU [sec]
1	3.2134	3.0474	3.1312	0.0183	0.0288	0.0204	4 x 256 <sup>3</sup>	11.19231	> 36.000
2	3.1112	3.0478	3.2231	0.0198	0.0269	0.0212	4 x 128 <sup>3</sup>	9.71301	4006.34
3	3.1507	3.0117	3.1160	0.0198	0.0269	0.0206	4 x 64 <sup>3</sup>	7.10882	397.21
4	3.1338	3.2669	3.3408	0.0176	0.0302	0.0227			

**Table 3.** Real patient data experient results. 256x256x106 voxels per volume (CT and US). (a) Standard deviations of translation and rotation parameters for random study. (b) Computational performance of simultaneous multi-modal registration for three and five volumes.

(a)							(b)	
Vol.	tx [mm]	ty [mm]	tz [mm]	$\alpha$ [rad]	$\beta$ [rad]	$\gamma$ [rad]	Volume Config.	Time [sec]
1	3.2829	3.3165	3.5544	0.0212	0.0331	0.0230	1CT,2US	333.93
2	3.2813	2.9114	2.9758	0.0173	0.0304	0.0232	1CT,4US	1076.57
3	3.2195	3.2878	3.8525	0.0176	0.0319	0.0259		
4	3.2345	2.8773	3.1018	0.0170	0.0311	0.0243		

The multi-modal mosaicing was evaluated qualitatively by visual inspection of alignment of major anatomical regions, e.g. organs interfaces, vasculature inside the liver. The improvement of alignment before and after registration is clearly visible (see Fig. 3), but has to be evaluated detailed and qualitatively in further studies. We validated the improved registration for problematic configurations by a testcase with two US volumes with very small overlap and one CT scan (see Fig. 2). Using only US intensity information the registration fails and breaks the link of the US volumes. Using our proposed mosaicing method, the US volumes are registered to each other and to the CT data (see Fig. 3).

## 4 Discussion

In this work we have presented a novel, multi-modal and multi-variate, approach to 3D US mosaicing. The additional information about the patient anatomy from CT allows mosaicing of US scan configurations with very small pair-wise volume overlap and reduces problems caused by US intensity variations. We demonstrated that direct, simultaneous mono and multi-modal 3DUS mosaicing works on clinical data. In future work we will conduct an extensive evaluation of the proposed method on phantom and real clinical datasets. Physician defined landmarks in CT and US data will be used to assess the quality and correctness of the method. Our work also addressed the current major limitation of direct, simultaneous, multi-variate image registration by implementing the most expensive computations on stream processors of computer graphics hardware. Our

vision is to extend it to a general multi-modality, multi-variate image registration framework. The objective is to perform multi-modal registration not only for data from a pair of modalities but to incorporate all information into the registration framework.

**Acknowledgments.** This work was partially supported by Siemens Corporate Research and funding from the European Committee within the passport project - <http://www.passport-liver.eu/>.

## References

1. Aiger, D., Cohen-Or, D.: Mosaicing ultrasonic volumes for visual simulation. *IEEE Computer Graphics and Applications* 20(2), 53–61 (2000)
2. Wachinger, C., Wein, W., Navab, N.: Registration strategies and similarity measures for three-dimensional ultrasound mosaicing. *Academic Radiology* 15, 1404–1415 (2008)
3. Gee, A., Treece, G., Prager, R., Cash, C., Berman, L.: Rapid registration for wide field of view freehand three-dimensional ultrasound. *IEEE transactions on medical imaging* 22(11), 1344–1357 (2003)
4. Poon, T., Rohling, R.: Three-dimensional extended field-of-view ultrasound. *Ultrasound in Medicine & Biology* 32(3), 357–369 (2006)
5. Ni, D., Qu, Y., Yang, X., Chui, Y.-P., Wong, T.-T., Ho, S.S.M., Heng, P.-A.: Volumetric Ultrasound Panorama Based on 3D SIFT. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) *MICCAI 2008, Part II*. LNCS, vol. 5242, pp. 52–60. Springer, Heidelberg (2008)
6. Wein, W., Brunke, S., Khamene, A., Callstrom, M.R., Navab, N.: Automatic ct-ultrasound registration for diagnostic imaging and image-guided intervention. *Medical Image Analysis* 12, 577–585 (2008)
7. Kutter, O., Shams, R., Navab, N.: Visualization and GPU-accelerated simulation of medical ultrasound from CT images. *Computer Methods and Programs in Biomedicine* 94(3), 250–266 (2009)
8. Karamalis, A., Wein, W., Kutter, O., Navab, N.: Fast hybrid freehand ultrasound volume reconstruction. In: *SPIE Medical Imaging, Orlando, Florida, USA (February 2009)*