

# Ultrasound based Respiratory Motion Compensation in the Abdomen

Wolfgang Wein, Jie-Zhi Cheng, and Ali Khamene

Imaging & Visualization Department  
Siemens Corporate Research, Princeton, NJ, USA  
{wolfgang.wein,ali.khamene}@siemens.com

**Abstract.** For image-guided interventions, the patient respiratory motion has to be considered when integrating pre-operative plans and imaging into the interventional suite. Ultrasound is particularly attractive for obtaining live information about anatomical changes, without adding further radiation dose to patient and physician such as fluoroscopy. We propose a strategy to both detect and compensate for respiratory movements in the liver, relying solely on real-time information from a tracked 2D ultrasound transducer, as well as one position sensor attached to the patient’s skin. Our method does not require a patient-specific 4D model of the respiration, which would be costly to establish and might not recover entire breathing variabilities and irregularities during the procedure. Appropriate transformation models, interpolation schemes and implementation details are discussed. By means of evaluation on liver data of four volunteers, we demonstrate the usefulness of our approach for ultrasound-based guidance of abdominal biopsies and ablations.

## 1 Introduction

### 1.1 Clinical Context

Treating cancer by destroying malignant tumor tissue deep inside the human body requires image-based localization and guidance. This can be achieved with maximum accuracy for static targets, e.g. in the case of stereotactic radiosurgery. If respiration and cardiac motion influence the tumor location dynamically, higher treatment margins have to be defined, possibly increasing trauma and compromising the effectiveness of the treatment. For radiotherapeutic treatment of lung cancer, gating approaches have started to be used clinically, where the treatment beam is only activated during a defined phase in the breathing cycle, as measured by a respiration sensor. Adjusting the treatment live to continuously update the tumor location, usually requires a patient-specific motion model derived from 4D pre-operative imaging, in combination with intra-operative imaging or measurement of surrogate signals. First clinical solutions are for adaptive radiotherapy are available that track tumor motion in X-Ray images, adjusting the treatment beam by means of a robot-mounted linear accelerator (Accuray Cyberknife).

In interventional oncology, ultrasound is the preferred intra-operative imaging modality for a large number of procedures, such as biopsies and thermal ablation in liver & kidney, laparoscopic liver surgery, or prostate biopsy. Navigation solutions are now available that utilize a tracked ultrasound transducer [1], and/or magnetically tracked needle tips (Traxtal Inc., CAS Innovations AG, etc.). Those systems share the ability to integrate pre-operative CT or MRI imaging into the live tracking workspace. Currently, the interventionalist has to use them in a stop-and-go manner, where he can only rely on correct alignment during the breathing phase in which the pre-op image was acquired. This is a rather similar approach as respiratory gating for radiotherapy.

## 1.2 Related Work

A number of research groups have proposed to pre-operatively create patient-specific models of the liver motion from 4D MRI acquisitions, e.g. [2]. Such a 4D motion model is registered in [3] to tracked 2D ultrasound images by optimizing over the single temporal parameter describing the deformation, in order to compensate the respiratory motion. A 2D+t MRI cine acquisition in a coronal plane is correlated with surrogate skin markers [4], using PCA based methods, to henceforth yield prediction of the internal motion from the surrogate signals alone. In [5] a 3D freehand laparoscopic ultrasound probe is used to establish a 4D motion model of a pig liver during the intervention, by acquiring a number of sagittal planes throughout the breathing cycle. Successively, the 2D deformations in those planes are estimated, and composed into a 4D motion model. A cross-correlation image similarity of successive frames serves as the signal indicating the respiratory phase.

All aforementioned techniques share the problem that establishing a 4D motion model (pre- or intra-operatively) is too time-consuming and costly to gain acceptance in the clinical practice, particularly for ‘lightweight’ procedures such as biopsies. Besides, they cannot reproduce breathing irregularities, which are very likely to occur on patients not fully sedated.

In our work, we seek to both detect and compensate the breathing-induced motion of a target within the liver, utilizing solely a tracked 2D ultrasound transducer and one position sensor attached to the patient’s skin.

## 2 Methods

### 2.1 General Approach

Before starting the procedure, the interventional radiologist acquires an ultrasound sweep over the whole liver during a breath-hold. This data henceforth serves as reference 3D information (denoted  $I_R$ ), and can be registered with a pre-operative CT or MRI scan, using e.g. methods described in [6]. During the procedure, the physician freely places the ultrasound transducer to guide needle insertion, and visualize & monitor the lesion.

Since the real-time ultrasound image  $I_{U_k}$  (at discrete time  $k$ ) is tracked, we can compensate for respiratory motion using a slice-to-volume registration approach:

$$\widetilde{\mathbf{a}}_k = \arg \max_{\mathbf{a}_k} \mathcal{M} \{ (I_R(\mathbf{M}_{\mathbf{a}_k}(\mathbf{p})), I_{U_k}(\mathbf{p})), \mathbf{p} \in \Omega_{I_U} \} \quad (1)$$

For every time  $k$ , an image similarity metric  $\mathcal{M}$  is maximized with respect to motion parameters  $\mathbf{a}_k$  of the mapping  $\mathbf{M}_{\mathbf{a}_k}$ , which maps the points  $\mathbf{p}$  from the 2D ultrasound image domain  $\Omega_{I_U}$  into the reference volume  $I_R$ . While abdominal motion is generally of deformable nature, a 2D ultrasound slice  $I_{U_k}$  does not contain sufficient information to recover a dense 3D deformation field. Since we are ultimately interested in tracking the displacement of a single clinical target within  $I_R$ , we model the motion as a global affine transformation. In terms of homogenous linear transformations, the mapping  $\mathbf{M}_{\mathbf{a}_k}$  is then composed as

$$\mathbf{M}_{\mathbf{a}_k}(\mathbf{p}) = \mathbf{A}_k^{-1} \mathbf{T}_R \mathbf{T}_k \mathbf{T}_C \mathbf{p}; \quad \mathbf{p} = (x, y, 0, 1)^T \quad (2)$$

where  $\mathbf{T}_C$  is the constant ultrasound to tracking sensor calibration,  $\mathbf{T}_k$  the tracking transformation at time  $k$ , and  $\mathbf{T}_R$  a constant registration transformation introduced for generality (for now assumed as identity transformation). Those are rigid transformation matrices, i.e.  $\mathbf{T}_R, \mathbf{T}_k, \mathbf{T}_C \in SE(3)$ . We are now seeking the parameters of  $\mathbf{A}_k$ , the 3D affine transformation expressing the deformation of the reference volume  $I_R$  onto the current ultrasound frame  $I_{U_k}$ .

## 2.2 Incorporating Surrogate Information

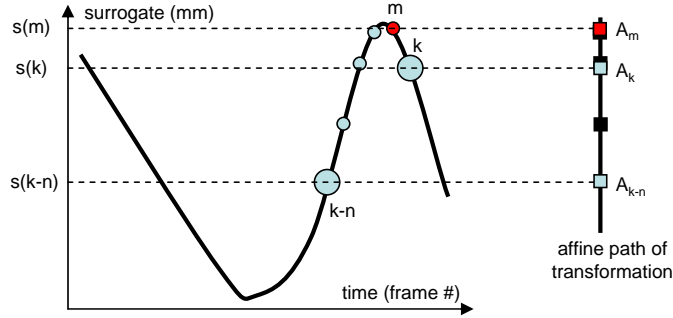
Optimizing the parameters of  $\mathbf{A}_k$  independently for every  $k$  results in a rather bumpy and unstable motion of the compensated volume  $I_R$ , since it solely depends on the current 2D US image, and the anatomical structures depicted therein. We therefore would like to constrain how  $\mathbf{A}_k$  can change over time, however without the strong assumption of a periodic motion, because we would like to be able to successfully recover breathing irregularities as well. From a position sensor attached to the patient's skin (close to the umbilicus), we extract a scalar surrogate measurement  $s(k)$ , representing the anterior-posterior translation. Instead of only using the current frame, we are registering the last  $n$  frames wrt. one set of affine parameters. We use the assumption that the  $n$  last affine transformations lie on a straight path, their actual location on it being derived from the surrogate signal (figure 1):

$$\mathbf{a}_{k-n+i} = \mathbf{a}_{k-n} + \frac{s(k-n+i) - s(k-n)}{s(m) - s(k-n)} (\mathbf{a}_k - \mathbf{a}_{k-n}); \quad i \in [1 \dots n] \quad (3)$$

where

$$m = \arg \max_i |s(i) - s(k-n)| \quad (4)$$

The transformation of frame  $m$ , whose surrogate signal is most distant from the oldest frame  $k-n$ , is optimized. The transformation of frame  $k-n$ , i.e. the  $n-1$ st frame in the past, is assumed constant, its image is not included in the similarity



**Fig. 1.** The surrogate signal is used for weighting along the transformation path.

measure computation. All transformations  $\mathbf{a}_{k-n+i}$  in between are interpolated - for now with a parameterization  $\mathbf{a}$  that supports addition (eq. 6 below). During inhaling and exhaling,  $m$  is the most current frame ( $m = k$ ). Around full in- and exhale,  $m$  is the frame closest to the extremum of the surrogate (see figure 1).

### 2.3 Representation of $\mathbf{A}_k$

We describe the sought affine transformation as

$$\mathbf{A}_k = \begin{pmatrix} \mathbf{HSR} & \mathbf{t} \\ \mathbf{0} & 1 \end{pmatrix}; \text{ with } \mathbf{H} = \begin{pmatrix} 1 & h_{xy} & h_{xz} \\ 0 & 1 & h_{yz} \\ 0 & 0 & 1 \end{pmatrix}; \quad \mathbf{S} = \begin{pmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & s_z \end{pmatrix} \quad (5)$$

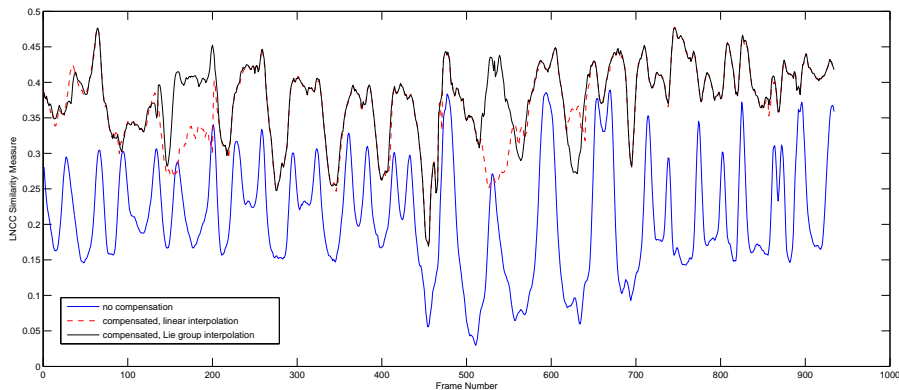
and  $\mathbf{R} = \mathbf{R}_z(\theta_z)\mathbf{R}_y(\theta_y)\mathbf{R}_x(\theta_x)$  an Euler angles rotation. The resulting parameterization is

$$\mathbf{a}_k = (t_x t_y t_z \theta_x \theta_y \theta_z s_x s_y s_z h_{xy} h_{xz} h_{yz})^T \quad (6)$$

Locally optimizing all 12 parameters for every frame is costly and potentially unstable (the scaling and shearing parameters easily bail out, depending on the amount of structures in the 2D ultrasound frame). We have used Principal Components Analysis (PCA) of the vector  $\mathbf{a}_k$ , created on ideal imaging data, in order to derive a better parameterization, as well as reduction of the degrees of freedom (DOF). Hence the first parameters of a vector  $\mathbf{a}'_k$  are optimized, such that  $\mathbf{a}_k = P\mathbf{a}'_k$ , where  $P$  is the  $12 \times 12$  matrix of PCA basis vectors. We obtain best results using only 3 DOF, representing the 3 most significant PCA modes and comprising 83% of the variance in the vector space of  $\mathbf{a}_k$ .

Another problem to be addressed is the interpolation of transformation matrices. Linearly weighting the actual parameters  $\mathbf{a}_k$  as in equation 3 is not accurate except for translation (e.g., an incremental motion applied twice would not yield twice the motion). Hence we use a Lie group based method for interpolation of  $\mathbf{A}_k$  [7]. It requires mapping  $\mathbf{A}$  in and out of the Lie group manifold's tangent space at  $\mathbf{A}_k$ , by means of matrix logarithm and exponential, respectively:

$$\mathbf{A}_{k-n+i} = \mathbf{A}_{k-n} \exp \left( \frac{s(k-n+i) - s(k-n)}{s(m) - s(k-n)} \log(\mathbf{A}_{k-n}^{-1} \mathbf{A}_k) \right) \quad (7)$$



**Fig. 2.** Similarity measure plot for a free breathing sequence.

## 2.4 Slice-to-Volume Registration

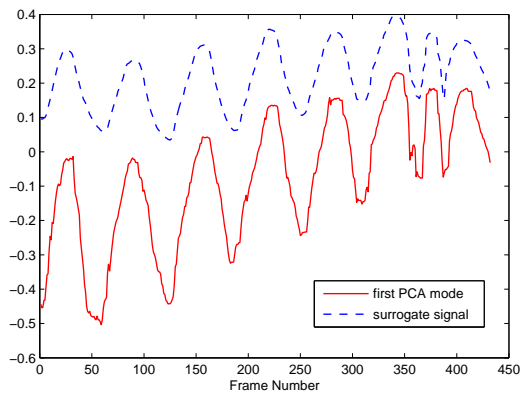
As the similarity measure ( $\mathcal{M}$  in eq. 1), Local Normalized Cross Correlation (LNCC) is used, implemented as recursive filter. While its computational effort is significantly higher than global metrics like Sum of Squared Differences (SSD) or Normalized Cross Correlation (NCC), we found it to be much more robust. This is due to its invariance to local brightness and contrast changes, induced by the orientation-dependency of ultrasound imaging. We did not consider ultrasound-specific similarity measures as in [8], because the involved resampling steps from 1) compounding of the inhale sweep into  $I_R$ , 2) slice extraction, and 3) down-sampling the 2D frame for registration, discard fine ultrasonic speckle noise features that are assumed by those measures.

The similarity measure is optimized wrt. the 3 transformation parameters  $\mathbf{a}'_k$ , using the Powell-Brent direction search technique [9]. Its specific advantage here is that the first Brent line optimization is executed on the most significant PCA mode, which recovers most of the mismatch and hence assures good performance and robustness. We restrict the optimization to a maximum of three iterations (each of them performing 3 line optimizations).

## 3 Results

The proposed technique was evaluated on four volunteers. Our tracked ultrasound setup comprised a Siemens Sequoia ultrasound machine (Siemens Ultrasound, Mountain View, CA) with abdominal curved-array probe, and a Micro-Bird magnetic tracking system (Ascension Technology Corp., Burlington VT). After acquiring an inhale liver sweep, recordings were made in free breathing, with the transducer moved arbitrarily to image any longitudinal and transversal planes of the liver. At the end of the recording, the volunteer would purposely breathe in an irregular manner.

We use  $n = 6$  frames for our algorithm, the acquisition rate is 10 Hz. Figure 2



**Fig. 3.** First PCA mode compared to surrogate signal (the latter is scaled and shifted to fit on the same plot).

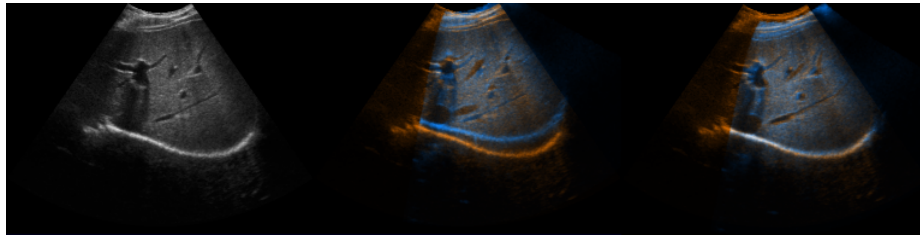
	CPU LNCC	CPU SAD	GPU SAD
time / evaluation	4.8 ms	1.9 ms	1.2 ms
evaluations / frame	43	38	41
time / frame	210ms	65ms	49ms

**Table 1.** Computation times. The CPU versions were executed on a Core 2 Duo 2.2GHz notebook, the GPU version on a NVidia Quadro FX 5500 graphics card.

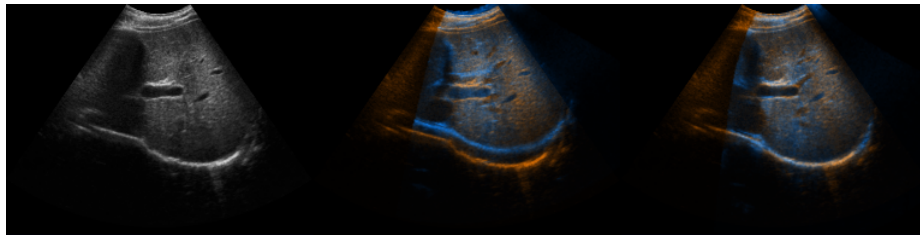
shows the similarity measure for an exemplary sequence. Our compensation has consistently higher similarity than the uncompensated version. Using the Lie group interpolation (eq. 7) yields better results compared to linear interpolation of the affine transformation parameters (eq. 3). The fact that the uncompensated signal sometimes has multiple peaks in inhale, can be explained by the patient’s inhaling exceeding the state where the reference sweep  $I_R$  was acquired.

Figure 3 depicts the first PCA motion parameter compared to the surrogate signal. One can see that the compensation smoothly correlates with the skin sensor signal. From visual assessment of the compensated sequences we concluded that the recovered motion accurately matched the live ultrasound feed. Besides, the student’s t-test was applied for the null hypothesis that the first PCA mode correlates with the surrogate signal. The resulting p-values on all four volunteers were larger than 0.50. This means there is no significant mis-match between those two signals (while the actual shape of the signals may well be different, since they represent internal vs. external motion). Figure 4 shows images for a number of breathing phases.

Our C++ implementation utilizes SSE2-accelerated slice extraction and OpenMP multi-threading (parallel extraction & comparison of slices). Computation times are shown in table 1. The local cross-correlation approach takes significantly



(a) mid-exhale



(b) mid-inhale

**Fig. 4.** Motion-compensated image pairs. Left is the live ultrasound frame, middle and right show color overlays of un-compensated and compensated ultrasound images. The hysteresis effect can be seen (different motion path during in- and exhaling).

more time than Sum of Absolute Differences (SAD). However, the numbers show that the motion compensation algorithm would run in real-time on a high-end quad-core workstation. We also compared the performance to a graphics processing unit (GPU) based version. Here, slice extraction was implemented as OpenGL 3D texturing, and SAD similarity measure computation in the fragment shader. For this particular algorithm, the GPU version did not boost the computation time much, because down-scaled ultrasound images with  $\approx 12000$  pixels were used to increase similarity measure smoothness. While the individual steps of slice extraction and fragment shader passes are extremely fast, a number of expensive render target switches (here with Frame Buffer Objects, FBOs) are required per slice (this holds also for recursive filtering to compute LNCC on the GPU, which we did not develop to date). For high resolution images, the GPU version would be an order of magnitude faster than the software implementation.

## 4 Conclusion

We have presented a novel technique for image-based compensation of internal liver motion, based solely on real-time tracked 2D ultrasound (also often denoted ‘3D freehand ultrasound’) imaging and a surrogate tracking sensor. It combines methods of 1) slice-to-volume registration with a local cross-correlation measure, 2) PCA-based reduction of the affine parameter space, 3) direction set optimiza-

tion, and 4) incorporation of a surrogate signal via Lie group interpolation of the  $n$  most recent motion updates. Real-time implementations of the algorithm were developed for CPU and GPU. Qualitative evaluation was done on four volunteers, resulting in visually convincing motion tracking even for highly irregular breathing.

We are currently acquiring data in a clinical setting on liver tumor patients. We will investigate the possibility to set up the 2D ultrasound imaging plane to monitor the lesion while at the same time providing enough anatomical clues (diaphragm, vasculature etc.) to allow our algorithm to precisely track tumor motion. Furthermore it is planned to evaluate the usefulness of the presented motion compensation for visualizing registered pre-operative data for interventional ablation guidance, comparing it to static registration (i.e. accurate alignment at only one time during the respiratory cycle, optional stop-and-go visualization). We believe that our approach will be useful for other clinical scenarios as well, particularly due to its main advantage over existing techniques, namely no need for 4D imaging and patient-specific motion modeling.

## References

1. Crocetti, L., Lencioni, R., DeBeni, S., See, T., Pina, C., Bartolozzi, C.: Targeting liver lesions for radiofrequency ablation: an experimental feasibility study using a CT-US fusion imaging system. *Investigative Radiology* **43** (January 2008) 33–39
2. Rohlfing, T., C.R. Maurer, J.: Modeling liver motion and deformation during the respiratory cycle using intensity-based nonrigid registration of gated MR images. *Medical Physics* **31** (2004) 427–432
3. Blackall, J., Penney, G., King, A., Hawkes, D.: Alignment of sparse freehand 3-D ultrasound with preoperative images of the liver using models of respiratory motion and deformation. *IEEE Trans. Med. Imag.* **24** (November 2005) 1405–1416
4. Khamene, A., Warzelhan, J., Vogt, S., Elgort, D., ChafdHotel, C., Duerk, J., Lewin, J., Wacker, F., Sauer, F.: Characterization of internal organ motion using skin marker positions. In: *MICCAI 2004 Proceedings*. Volume 3217 of *Lecture Notes in Computer Science*. (2004) 526–533
5. Nakamoto, M., Hirayama, H., Sato, Y., Konishi, K., Kakeji, Y., Hashizume, M., Tamura, S.: Recovery of respiratory motion and deformation of the liver using laparoscopic freehand 3D ultrasound system. *Medical Image Analysis* **11** (November 2007) 429–442
6. Wein, W., Khamene, A., Clevert, D., Kutter, O., Navab, N.: Simulation and fully automatic multimodal registration of medical ultrasound. In: *MICCAI 2007 Proceedings*. Volume 4791 of *Lecture Notes in Computer Science*., Springer (October 2007) 136–143
7. Subbarao, R., Meer, P.: Nonlinear mean shift for clustering over analytic manifolds. In: *CVPR Proceedings*. (2006) 1168–1175
8. Cohen, B., Dinstein, I.: New maximum likelihood motion estimation schemes for noisy ultrasound images. *Pattern Recognition* **35** (2002) 455–463
9. Press, W., and W.T. Vetterling, S.T., Flannery, B.: *Numerical Recipes in C, Second Edition*. CRC Press, Inc. (1992)