

Real Time Image-Based Tracking of 4D Ultrasound Data

Ola Kristoffer Øye^{1,2}, Wolfgang Wein³,
Dag Magne Ulvang², Knut Matre⁴, and Ivan Viola¹

¹ Department of Informatics, University of Bergen, Norway

² Christian Michelsen Research, Bergen, Norway

`olak@cmr.no`

³ Chair for Computer Aided Medical Procedures (CAMP),
Technische Universität München, Germany

⁴ Institute of Medicine, University of Bergen, Norway

Abstract. We propose a methodology to perform real time image-based tracking on streaming 4D ultrasound data, using image registration to deduce the positioning of each ultrasound frame in a global coordinate system. Our method provides an alternative approach to traditional external tracking devices used for tracking probe movements. We compare the performance of our method against magnetic tracking on phantom and liver data, and show that our method is able to provide results in agreement with magnetic tracking.

Keywords: Ultrasound, image registration, real time image processing.

1 Introduction

Medical ultrasound is a flexible, non-invasive, hand-held imaging modality used in a wide range of clinical applications. There is increased interest to assign precise 3D location information to ultrasound data, in order to enable reproducible results and diagnostic processes similar to how e.g. CT and MRI data are used. Such approaches often suffer from motion of the examined subject due to body movement, respiration and heart beat. Navigation in neurosurgical treatments is utilizing high-precision optical tracking systems by fixing optical markers to the cranium, surgical instrument, and ultrasound probe so that their mutual positions are known and involuntary movements are avoided. In the scenario of abdominal ultrasound, large organs in general do not fit into a single scan sector and they exhibit significant shifts due to respiration movements. In this scenario, external positional magnetic tracking is employed to measure position and rotation of the US probe in 3D space. This does not require line-of-sight, but suffers from other severe drawbacks. Most importantly, it tracks the position of the probe and not of the subject of examination, therefore assumptions have to be made that neither the magnetic transmitter nor the subject did move. This assumption is carried out by asking the patient to perform shallow breathing, or to stop breathing for a limited time period within the same phase of

a respiration cycle. These workarounds reduce to some degree positioning errors, but are sensitive to patient movement and unreliable when high precision is needed. Also, any presence of ferromagnetic objects deforms the magnetic field and makes the measurements imprecise. Thus, reliable positioning technology is urgently needed for abdominal US examinations. Our concept addresses exactly this challenge. First, we propose a novel tracking approach, where the positioning information is obtained internally, as a result of image registration of streaming US 3D data. Without an external reference, systematic errors such as drift influence the precision. As a second contribution, we therefore present effective strategies to minimize the drift. Third, registration as a tracking concept is carried out as a fast GPU implementation that converges sufficiently fast so that the probe can be tracked in real-time.

Related Work: A method for intensity-based registration of 3D US of the prostate is presented in [1]. In [2], a variational method is proposed to register a single pair of 3D US volumes of the liver while considering soft tissue deformations. [3] develops strategies for the mosaicing of multiple US volumes, including the use of multi-variate similarity measures. Designated similarity measures to deal with ultrasonic noise statistics have been developed as well, see e.g. [4]. All those works have however focused on the methodology as opposed to real-time computation. Work in [5] presents a fast GPU implementation and adds the registration to CT as a global map guiding the local US volumes. In [6], real-time 2D US with magnetic tracking is used to establish abdominal 4D US covering respiratory motion. Our new contribution is a designated approach for fully replacing an external tracking system using real-time registration of 4D US.

2 Methodology

Our image based tracking consists of two main steps (Fig. 1). First, a sweep over the organ of interest is performed. Using *dead reckoning tracking*, the individual frames from the sweep are mosaiced into a larger *compound volume*. Second, the compound volume is used as the reference for tracking streaming US data, using *reference volume tracking*. We have developed a solution that streams data real time from a GE Vivid E9 scanner over ethernet to any computer running our software, providing the registration framework with a stream of US volumes.

“Dead Reckoning” Image Based Tracking: In order to build the compound volume from streaming US data without an external tracking reference, we deduce the position of the current image from the position of the previous images, a principle known as “dead reckoning” from navigation. The most basic approach would be to use pairwise (PW) image registration to find the relative transformation T between the current image and the previous image in the stream. For each new image, we could use the product of all previous T s to find the position of this image in a common coordinate system. This will however propagate the error from every PW registration to all following frames, leading to a significant drift over time. We therefore extend the classical PW image registration to multi-frame registration: Instead of registering a moving image to a single

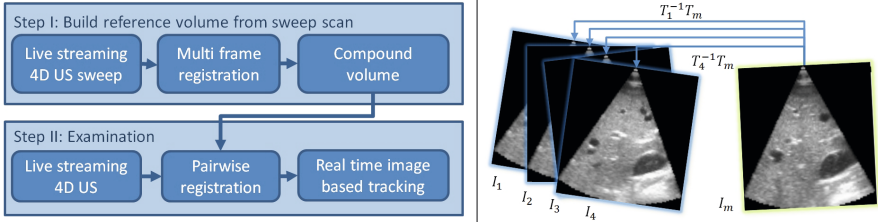


Fig. 1. Image based tracking workflow (left) and multiframe registration (right)

fixed image, we register our image to n fixed images in a global coordinate system (Fig. 1). For the live streaming case, we use the n most recent images. We use the sum of squared differences (SSD) as our image similarity measure. Specific US similarity measures have been developed [4], but their practical advantage has never been conclusively shown for clinical applications. Besides, the image noise statistics assumed for such methods are generally manipulated further down in the US processing chain (i.e. time-gain compensation, log-compression, scan-conversion etc.). Since we are registering neighbouring frames, we can assume that the anatomical regions have echoes of similar intensity due to only small changes in probe orientation. Given a linear transformation T , a fixed image I_f and a moving image I_m , we define the similarity between two images as

$$d(I_f, I_m; T) = \frac{\sum_x [I_f(x) - I_m(Tx)]^2}{N} + \begin{cases} 0 & \text{if } N \geq N_0 \\ \beta(1 - N/N_0) & \text{if } N < N_0 \end{cases} \quad (1)$$

where the sum is taken over all voxels x in the image and N is the number of overlapping voxels. The last term is a penalty function to avoid artificial low overlap minima, added for overlaps less than N_0 (typically 20% of total number of voxels). β is a constant penalty weight. Only voxels with an intensity over a certain threshold are included in the SSD calculation to avoid taking shadow regions into account. Based on (1), we construct a similarity measure D for a moving image vs. any number of fixed images in a common frame of reference:

$$D(\mathbf{I}^{(n)}, \mathbf{T}^{(n)}, I_m, T_m) = \sum_{i=0}^n \alpha_i [d(I_i, I_m, T_i^{-1}T_m)] \quad (2)$$

where \mathbf{I} and \mathbf{T} are vectors of size n containing the set of fixed images and their respective global transformations. The factor α_i is the fraction of total overlap for frame i , giving more overlapping frames stronger contribution. For $n = 1$, the equation reduces to PW image similarity. We allow for two types of linear transformations: rigid transformation with 6 degrees of freedom (DOF) and affine transformation without scaling with 9 DOF. We use the rigid transformation for dead reckoning tracking for compounding, assuming that the object does not deform under the compound sweep. The registration problem can be stated as:

$$\min_x D(\mathbf{I}^{(n)}, \mathbf{T}^{(n)}, I_m, T_m(x)) \quad (3)$$

where x is the vector containing the parameters specifying the rigid (affine) transformation with 6 (9) DOF. We use the nonlinear Nelder-Mead optimizer to minimize Eq. (3). The outcome of the registration of image I_m is the global transformation T_m , defining the position of the image in a common coordinate system with all previous registered frames, essentially providing the path of the ultrasound scan relative to the scanned object.

Compound Reference Volume Construction: Using the dead reckoning tracking, we can resample each frame with its T_m into a larger compound volume. The value of each individual voxel of the compound volume is updated for each new resampled frame by averaging over the current and all previous contributions to each particular voxel, as described in [7], appendix B.

Reference Volume Image Based Tracking: We can now use the compound volume as a fixed image, and streamed volumes as moving images to track probe movement, avoiding accumulated drift. This is the tracking mode that is to be used in an examination phase. We use our registration approach with $n = 1$ and with affine transformation. This is to allow for some deformation due to breathing and patient movement. Since we track wrt. a reference anatomical map instead of an external reference, the method is independent of patient movement and up to a certain level organ deformations. When starting the registration, an initial search for a good fit is performed using a large step size in the optimizer. Once a good fit is found, for each new frame, we seed the registration with the position of the previous frame. If the tracking fails, the user reinitializes the tracking.

Implementation: For real time performance, we implement registration and compounding on the GPU using OpenCL. The pipeline starts with the streaming providing the most recent US volume in GPU memory. The similarity (1) is computed on the GPU: given T , I_f and I_m , the calculation of a squared difference image (SDI) is fully parallizable. A parallel sum reduction is performed on the SDI, producing the scalar SSD, which is transferred to the CPU where the optimizer is running. For the compounding, we use a kernel that given T resamples the streaming volume to the compound volume for each voxel in parallel.

3 Results

We evaluate the performance of our proposed method on phantom and liver scans, using a GE Vivid E9 scanner with a 4V matrix array 3D probe with a typical framerate of 10 FPS. We compare image based tracking to magnetic tracking by attaching a magnetic tracker sensor (Flock of Birds) to the probe. This system is today used in commercial systems for compounding and landmark positioning ('GPS'). As the magnetic sensor can not be placed exactly on the transducer element, we need to calibrate the offset T_C between the magnetic sensor and the image coordinate system. We use the procedure in [7], appendix

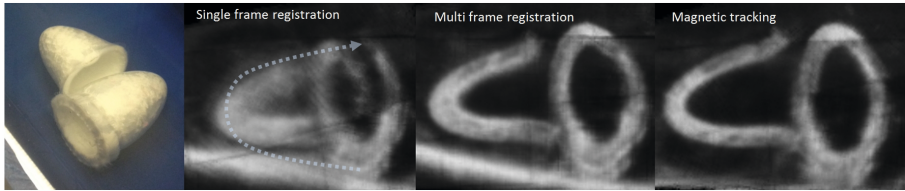


Fig. 2. Phantom used for evaluation (left). Compound volumes produced from image based tracking and magnetic tracking (top view).

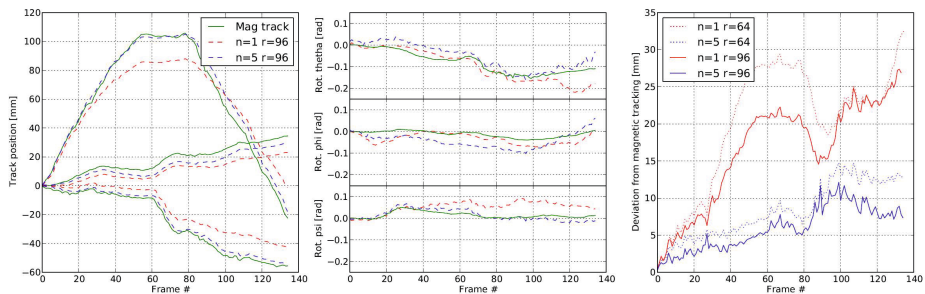


Fig. 3. Comparison of magnetic and image based tracking for sweep over phantom. Left: Tracked position of a point at the centre of the US volume, shown for all three coordinate components individually. Mid: Rotational components. Right: Discrepancy between magnetic and image based tracking for a point in the centre of the US volume.

B: Three volumes are acquired with external tracking (T_1, T_2, T_3). The relative T s between the volumes are found using image registration. Using optimization, we find the rigid tracker offset T_C that minimizes the difference in relative movement as measured by the magnetic tracker and by the image registration.

Evaluation on Phantom: For our evaluation we use two heart phantoms in a water tank (Fig. 2). We calibrate T_C using the phantom, and from the variation of T_C over multiple calibrations, we find that the calibration has an error of the order of mms and a few degrees of rotation. The error is a combination of the magnetic tracking positioning error and the error in the image registration.

We start by performing a sweep over the phantom to build a compound volume. Fig. 3 shows the reconstructed probe path and the discrepancy between the calibrated magnetic tracking and our image based tracking for $n=1, 5$ and volume number of voxels $r=64^3, 96^3$ (in our streaming we can vary the resolution of the scan converted US volume). The path of the sweep is indicated by the dotted arrow in Fig. 2. We see that for $n = 1$ (PW registration), there is a quite large drift between the image based and the magnetic tracking for both resolutions. By increasing n , we see that the discrepancy is significantly reduced. Further increases in n or r did not improve the results further. Using the tracking information, we build compound volumes for each of the cases. Slices from

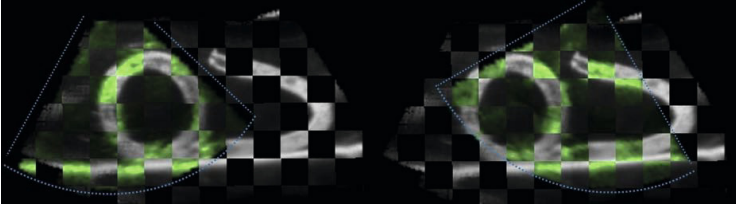


Fig. 4. Slice view of two time steps of image based tracking using compound volume as tracking reference. Streaming data in green, field of view indicated by dotted lines.

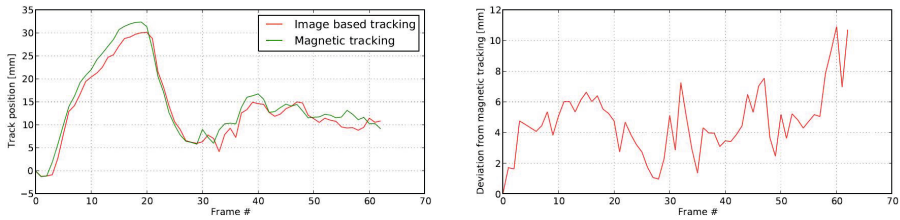


Fig. 5. Image based tracking for phantom scan using compound volume reference

the resulting volumes are shown in Fig. 2. We see that the PW registration has a quite large error as expected from Fig. 3, leading to a fuzzy and misaligned compound image. For the $n = 5$ case however, we see that we produce a compound volume with no visible drift. We actually get a better result than for the magnetic tracking, where we see that there is a shift in the right part of the image towards the end of the sweep. This can be explained by imperfections in the magnetic tracking calibration, and shows that magnetic tracking does not provide an absolute ground truth in our case. We then evaluate the *reference volume tracking* using the $n=5$ compound volume as a fixed image. To test the robustness of the method, we use a different probe path and a different FOV on the scanner from what was used when building the compound volume. Results are shown in Fig. 4, and we see that there is a good match between the compound volume and the streamed data. In Fig. 5 left, we compare with the magnetic tracking, plotting the distance between the centre of the tracked current frame and the initial frame. The right plot shows the distance between the centre point as tracked by magnetic and image based tracking. We see a discrepancy of the same magnitude as seen in Fig. 3. This is as expected, since we are using the compound volume from this scan as our tracking reference. We are also able to make a visual qualitative assessment that the image based tracking is able to follow the "anatomical" features of the phantom in the reference volume. As the $n=5$ compound volume was of higher quality than the one produced using magnetic tracking, this indicates that the image based tracking in this case provides a tracking of higher *anatomical* precision than the magnetic tracking.

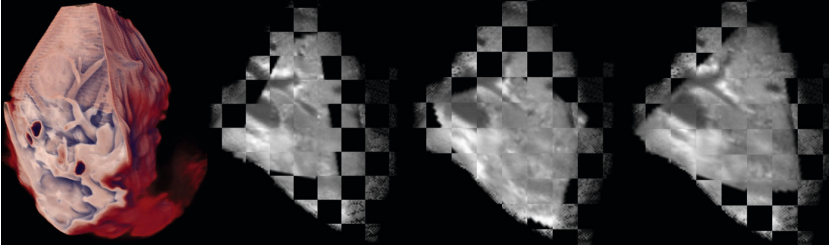


Fig. 6. Volume rendering of compound liver volume (left), three examples of streaming data with varying field-of-view registered to a reference compound volume for tracking

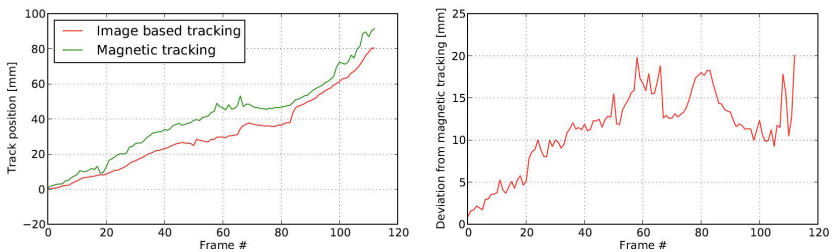


Fig. 7. Comparison of image based and magnetic tracking for a liver scan

Evaluation on Liver Scans: Having seen that we are able to build a compound volume and perform image based tracking on a phantom, we move to liver scans of a healthy volunteer (HV). We perform a sweep scan to build a compound volume while the HV is performing a deep breath hold. We avoid putting heavy pressure on the probe to minimize organ deformation. The subject is then allowed to breathe freely for a while, and then asked to perform a breath hold again. A new scan is then tracked wrt. to the compound volume, and we see from Fig. 6 that we are able to track the image in the anatomical reference of the compound volume using affine registration. We are using a FOV different from what was used for building the compound volume. Fig. 7 shows a comparison of the path reconstructed from the image based and magnetic tracking. They follow the same shape, but some discrepancy is expected due to the fact that the image based tracking registers into the anatomical reference, thus following an *anatomical path*, while the magnetic tracking simply follows the spatial path of the sensor. Since the HV had a period of breathing between the scans, there will inevitably be some movement and deformation of the liver compared to the initial sweep. Using affine registration, we are however able to handle this. This is an important point, since anatomical markers such as "GPS points" placed using magnetic tracking will be invalidated after patient movement, while our method is not affected by global patient movement, and able to handle some organ deformation.

Performance: We have tested our method on an Nvidia GTX 580, and show framerates for image based tracking for varying n and r in the table below. We are using (1, 64) for the reference volume tracking results shown. We have also compared our results with a corresponding CPU based setup using C++ ITK. We find that PW reg. at $r = 64(96)$ takes ~ 10 (35)s per frame on an Intel i5 2.5 GHz, making our GPU implementation faster by two orders of magnitude.

(n, r)	(1, 64)	(5, 64)	(1, 96)	(5, 96)
FPS	8.2	2.5	4.5	1.2

4 Conclusion

We have presented an approach for real time image based tracking, and shown that we achieve results that are in agreement with magnetic tracking, currently the industry standard for tracking in abdominal US. The major advantage of our method is that we are tracking the location of the probe wrt. anatomy instead of an external tracking reference. This means that the tracking precision of our method is not affected by global patient movement. We are able to track the probe relative to the organ also under some deformation, in contrast to magnetic tracking which assumes an entirely static organ. For further work on evaluation, optical tracking or a phantom with fiducial features should be used. We also intend to improve our algorithms for deformable abdominal motion, which may eventually be used to create large-scale 4D US over the entire respiratory cycle.

References

1. Baumann, M., Mozer, P., Daanen, V., Troccaz, J.: Towards 3D Ultrasound Image Based Soft Tissue Tracking: A Transrectal Ultrasound Prostate Image Alignment System. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part II. LNCS, vol. 4792, pp. 26–33. Springer, Heidelberg (2007)
2. Zikic, D., Wein, W., Khamene, A., Clevert, D.-A., Navab, N.: Fast Deformable Registration of 3D-Ultrasound Data Using a Variational Approach. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006, Part I. LNCS, vol. 4190, pp. 915–923. Springer, Heidelberg (2006)
3. Wachinger, C., Wein, W., Navab, N.: Three-Dimensional Ultrasound Mosaicing. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part II. LNCS, vol. 4792, pp. 327–335. Springer, Heidelberg (2007)
4. Wachinger, C., Navab, N.: Ultrasound Specific Similarity Measures for Three-Dimensional Mosaicing. In: Proc. SPIE Medical Imaging 2008, San Diego, USA (2008)
5. Kutter, O., Wein, W., Navab, N.: Multi-modal Registration Based Ultrasound Mosaicing. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009, Part I. LNCS, vol. 5761, pp. 763–770. Springer, Heidelberg (2009)
6. Nakamoto, M., et al.: Recovery of respiratory motion and deformation of the liver using laparoscopic freehand 3D ultrasound system. Medical Image Analysis 11(5), 429–442 (2007)
7. Wein, W.: Multimodal Integration of Medical Ultrasound for Treatment Planning and Interventions. PhD dissertation, Fakultät für Informatik, Technische Universität München (2007), <http://mediatum2.ub.tum.de/doc/620979/620979.pdf>