A 3D-Elastography-Guided System for Laparoscopic Partial Nephrectomies

Philipp J Stolka*(1), Matthias Keil(2), Georgios Sakas(2), Eliott McVeigh(1), Mohamad E Allaf (3), Russell H Taylor(1), Emad M Boctor(3)
(1) – Johns Hopkins University, Baltimore, MD/USA
(2) – Fraunhofer Institut für Graphische Datenverarbeitung, Darmstadt, Germany
(3) – Johns Hopkins Medical Institutions, Baltimore, MD/USA

ABSTRACT

We present an image-guided intervention system based on tracked 3D elasticity imaging (EI) to provide a novel interventional modality for registration with pre-operative CT. The system can be integrated in both laparoscopic and robotic partial nephrectomies scenarios, where this new use of EI makes exact intra-operative execution of pre-operative planning possible. Quick acquisition and registration of 3D-B-Mode and 3D-EI volume data allows intra-operative registration with CT and thus with pre-defined target and critical regions (e.g. tumors and vasculature). Their real-time location information is then overlaid onto a tracked endoscopic video stream to help the surgeon avoid vessel damage and still completely resect tumors including safety boundaries.

The presented system promises to increase the success rate for partial nephrectomies and potentially for a wide range of other laparoscopic and robotic soft tissue interventions. This is enabled by the three components of robust real-time elastography, fast 3D-EI/CT registration, and intra-operative tracking. With high quality, robust strain imaging (through a combination of parallelized 2D-EI, optimal frame pair selection, and optimized palpation motions), kidney tumors that were previously unregistrable or sometimes even considered isoechoic with conventional B-mode ultrasound can now be imaged reliably in interventional settings. Furthermore, this allows the transformation of planning CT data of kidney ROIs to the intra-operative setting with a markerless mutual-information-based registration, using EM sensors for intra-operative motion tracking.

Overall, we present a complete procedure and its development, including new phantom models – both *ex vivo* and synthetic – to validate image-guided technology and training, tracked elasticity imaging, real-time EI frame selection, registration of CT with EI, and finally a real-time, distributed software architecture. Together, the system allows the surgeon to concentrate on intervention completion with less time pressure.

Keywords: ultrasound, elastography, navigation, partial nephrectomy, registration, distributed software

1. INTRODUCTION

1.1 Motivation

Laparoscopic interventions are notoriously difficult for surgeons because of the limited field of view and limited dexterity associated with transcutaneous access. Laparoscopic partial nephrectomies – the standard of care treatment for a range of kidney cancers – compound these difficulties when sub-surface tumors are to be resected without the aid of visible landmarks for orientation. The current state-of-the-art approach of planning the intervention on CT data with only endoscopic intra-operative navigation is therefore error-prone because of limited intra-operative feature visibility. A major difficulty of these procedures is the time pressure the surgeon is facing after clamping the renal artery. The entire procedure of tumor localization, resection, stopping blood loss, and closure of the kidney has to be completed within 30 minutes. The complex internal renal anatomy (Figure 1) complicates this up to the point where indicated partial nephrectomies may become impossible. Therefore, advances in intra-operative imaging, registration with pre-operative plans, and navigation promise to substantially increase safety and applicability of these interventions.

Medical Imaging 2010: Visualization, Image-Guided Procedures, and Modeling, edited by Kenneth H. Wong, Michael I. Miga, Proc. of SPIE Vol. 7625, 76251I · © 2010 SPIE · CCC code: 1605-7422/10/$18 · doi: 10.1117/12.844589

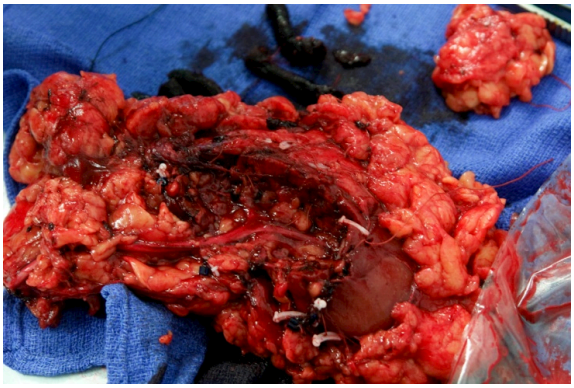

Figure 1: Excised kidney (front) after aborted partial nephrectomy. Note its complex internal anatomy with blood vessels close to the resected tumor (top right).

1.2 Approach

We propose a tracked laparoscopic system (Figure 2) for ultrasound-based 3D strain imaging (elasticity imaging/EI) to provide a novel modality for registration with pre-operative CT. Quick acquisition and registration of EI volume data allows intra-operative registration with CT and, by extension, with pre-defined locations of target and critical regions (e.g. tumors and vasculature). Dynamic registration or tracking is maintained by tracking the positions of the camera and the target regions using an electromagnetic (EM) localizer. The real-time location information of the target regions is then overlaid onto the tracked endoscopic video stream (the simple video amounts to the current standard procedure) to help the surgeon avoid vessel damage and still completely resect the tumor including safety boundaries.

2. SYSTEM

2.1 Methods

This approach is facilitated through real-time elastography, fast 3D-EI/CT registration, and intra-operative tracking. While real-time elastography modules do exist for commercial ultrasound (US) systems, they are extremely sensitive with respect to suboptimal palpation motion. It is difficult for sonographers to generate proper palpation motions (axial motion at the right speed) with conventional handheld probes; this is exacerbated by the use of laparoscopic US probes (*lapUS*) at the distal end of long tools. We therefore propose a combination of methods to improve the robustness of laparoscopic EI (*lapEI*).

First, the strain computation process itself is optimized. Unlike current approaches that mostly rely on cross-correlation or phase-zero methods, dynamic programming algorithms to compute the underlying tissue displacement maps exhibit fewer dropouts and naturally include displacement continuity [5] (Figure 4). Second, extending the conventional 1D (A-line) displacement estimation to a 2D method makes the process more robust under lateral (non-axial) motions. As this process can be partially parallelized, it has been ported to a GPU-based implementation to increase the achievable frame rate [1]. Finally, as the linear lapUS probe is electromagnetically tracked to allow 3D volume reconstruction, this information can be used to select optimal frame pairs for strain computation based on their relative poses [2]. This significantly reduces the amount of low-quality EI frames being generated (Figure 5), and improves the quality of the reconstructed volume.

The second key step in the proposed system is registration of the intra-operative situation with the pre-operative CT and its associated intervention plan. With the availability of EI, being a simple imaging modality sensitive to relevant features, previously segmented regions (the "plan") can be associated with the patient anatomy via registration of 3D-EI and CT. This registration step, however, is hampered by a variety of difficulties related to ultrasound. Apart from the need for tracking the probe, the field-of-view as well as the amount of discernible features are limited. As the region-based stiffness values of EI data are related to CT density information, an information-based metric (Mattes Mutual Information in this case) is used for registration [4].

Finally, these static registrations are maintained dynamically by associating an EM sensor with each component. With the segmentation plan rigidly associated with the underlying CT, which in turn is registered to the intra-operative tracked 3D-EI, the pre-operative plan regions are then present in tracking coordinates. One EM sensor is inserted beforehand

into the tumor itself under US guidance to ensure proper association of the tumor region of interest (ROI) with the actual site while the kidney is moved around during the intervention. Previous work in our group has augmented an endoscope (a standard component in laparoscopic surgery) with calibration and EM tracking, so the ROIs – represented as 3D surface objects – can then be overlaid onto the video stream in a real-time augmented reality (AR) setup. This serves as immediate feedback to the surgeon as to where he is allowed to cut into the tissue without damaging underlying structures or leaving behind tumor tissue that was to be resected.

All of the described steps were validated using a variety of realistic kidney phantom models, both physiological and synthetic, specifically aimed to mimic actual kidneys, to demonstrate feasibility of the proposed system. More detailed descriptions will be given in the following sections.

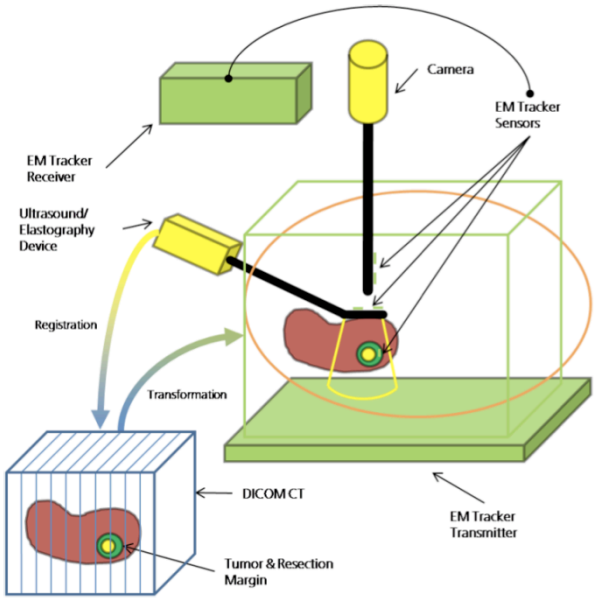

Figure 2: Proposed system components, comprising intra-operative situation (center), EM tracker (top left), endoscopic camera and ultrasound probe (around center), and pre-operative CT with planning data (bottom left).

2.2 Kidney Phantoms

The system has been developed and validated on both physiological porcine kidney phantoms as well as on synthetic ones. The challenges faced by surgeons in the actual intervention are closely tied to the visibility of the target and critical regions (i.e. tumor lesions as well as blood and collective vasculature) within the kidney. In pre-operative CT, both kinds of regions are discernible, which allows planning of the intervention. However, while the tumors may be palpable intra-operatively, they are often isoechoic, i.e. invisible under standard B-mode ultrasound. Furthermore, manual palpation is all but impossible in laparoscopic interventions. Therefore, the pre-operative plan can be transferred into the OR situation only approximately, with potentially catastrophic consequences for the operation outcome, where tumor excision in a partial nephrectomy can turn into total resection. Strain imaging can detect lesions of different stiffness deep within the tissue, which makes tumor detection and delineation possible.

In freshly-excised physiological phantoms (, flushed with saline solution), both thermal ablation and material injection create lesion areas with properties similar to real tumors. Contrast material injection resulted in CT images comparable to contrast-enhanced human kidney CT scans in different phases (cf. [4] for more details). With these phantoms exhibiting a high variability with respect to time-to-decomposition, a decision to transfer these results to synthetic phantoms was made. Based on our experience with kidney and lesion phantoms, we designed novel “Hopkins kidney phantom” prototypes with inserted lesions of varying stiffness and B-mode visibility. These were produced by CIRS Inc. (Norfolk, VA) and proved to be similar to physiological ones in all relevant aspects (cf. Figure 10 for those Hopkins phantoms in use; Figure 6 for EI from a different synthetic phantom). Details for both kinds of phantoms will be presented in a separate publication.

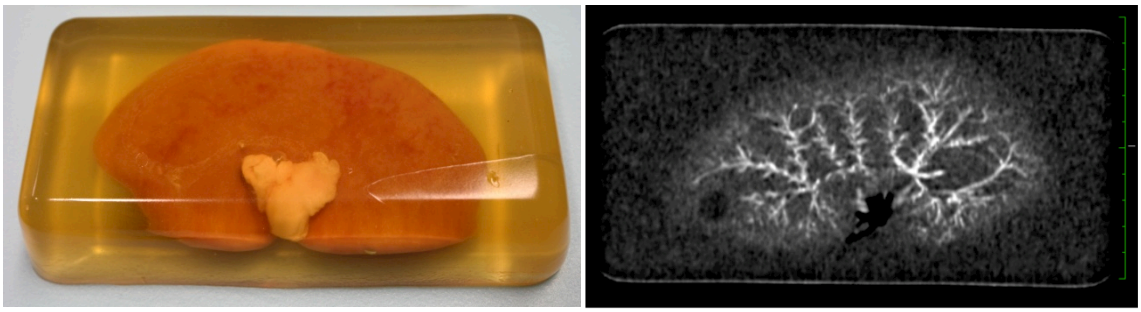

Figure 3: Physiological porcine kidney phantom. Embedded in gel (left), reproducible strain imaging of the contours as well as of internal structures (right; CT with contrast material in arterial-phase analogue and natural tumor) is possible.

2.3 Elasticity Imaging

In the context of the presented system, elasticity imaging (EI) makes it possible to visualize a class of renal tumors that are invisible in B-Mode due to their isoechoicity. In particular for laparoscopic ultrasound with a small field-of-view, the depiction of internal structures including tumors is important for successful registration with CT. Furthermore, EI images are by nature speckle-free and, even more importantly, "tissue-region-based" in a fashion similar to CT (unlike ultrasound B-Mode), which makes it possible to register them with CT without any further preprocessing.

2.3.1 Strain Computation Algorithms

The computation of static elasticity (or strain) images from ultrasound requires the availability of RF data. While performing subtle axial palpation motions, pairs of RF frames ("compressed" and "uncompressed") are subjected to an EI algorithm computing displacement maps first, followed by differentiation to generate strain maps.

To achieve a frame rate high enough for 2D real-time EI, two algorithms (normalized cross-correlation/*NCC* and dynamic programming/*DP*) were modified for parallel GPU-based computation. Because NCC is a local method in the sense that the computation of displacement for one sample point does not depend on the neighboring ones, it can be parallelized in a straightforward fashion along single axial RF lines. DP on the other hand relies on a recursive top-to-bottom estimation of displacements, where each displacement sample depends on the previous one. Therefore, DP cannot be easily parallelized along single lines; however, it offers higher axial resolution.

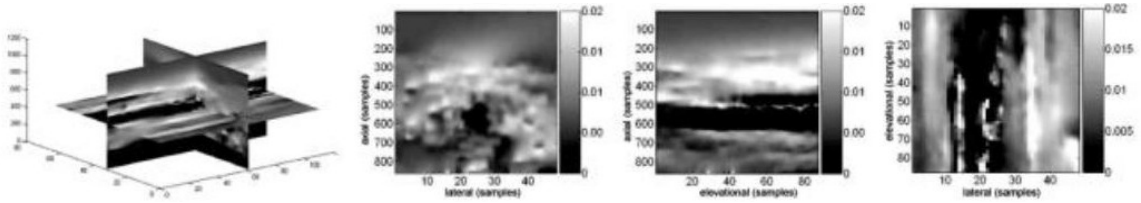

Figure 4: 3D EI volume (computed with a dynamic programming algorithm) for ablation monitoring, acquired with a 3D wobbler US probe (taken from [Rivaz08]).

2.3.2 Approaches for Quality Improvement

In the context of the presented system, three different approaches to improve the quality of the resulting EI stream were developed and investigated. Based on the two continuous sequential streams of input RF frames and their tracking positions, it possible to observe and modify the strain computation data both on the input and on the output sides.

Before feeding the raw streams to the EI module, optimal RF frame pairs can be selected based on their relative poses [2] (Figure 5). This exploits that evaluation of a large set of possible combinations based on their 6-DoF poses is very cheap compared to the actual computation of the corresponding EI frames. For every two poses P_1 and P_2 available as homogeneous 4×4 matrices, their relative displacement can be easily computed as $D_{12} = P_1 \cdot P_2^{-1}$. For an existing set of n RF frames and their poses, the possible combinations form a set of size $f(n) = \sum_{i=1}^{n}(i-1) = \sum_{i=1}^{n-1} i = \frac{n(n-1)}{2}$, where for each new frame/pose the oldest one is dropped and n–1 new pose combinations need to be computed. If the relative

displacement of one such combination consists of mainly axial (and possibly lateral) translations, the resulting strain images may have a high quality, while elevational or rotational displacement components decrease the quality due to speckle decorrelation that cannot be picked up with 2D (or generally slice-based) strain computation algorithms.

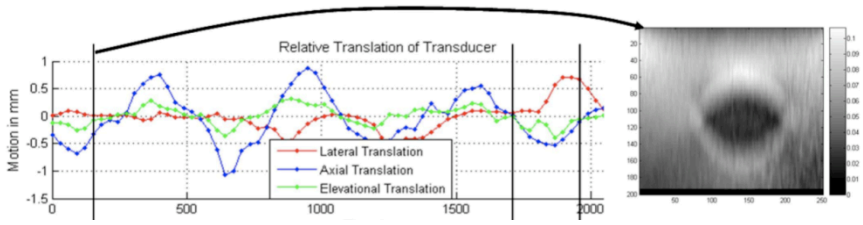

Figure 5: Selection of optimal frames for EI computation based on relative pose, using a 6-DoF EM tracker for position information (from [2]). Only frame pairs with mainly axial translations between the frames result in high-quality EI frames.

A second approach to improve the average quality of the strain image stream is to filter out low-quality frames after computing them. One possibility that was investigated is based on the NCC algorithm described in the previous paragraph. For each displacement sample computation (over search regions along one RF line), the maximum cross-correlation value is stored, yielding a cross-correlation map. The average over this map (average cross-correlation or ACC) can serve as a useful indicator of the average quality of that frame. Regions of the frame corrupted by noise will contribute with low values, while continuous matching regions increase it. Dropping frames below a threshold value of around tACC = 0.7 effectively filters out noisy images.

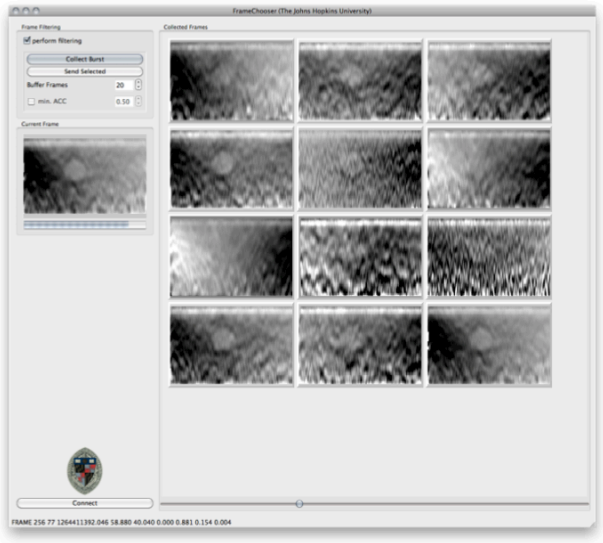

Figure 6: Frame display module, showing a collected burst of handheld 2D strain images with one spherical lesion, allowing semi-automatic selection of suitable frames for further processing. The real-time stream was computed using the NCC algorithm. Notice the quality indicator underneath the current frame shown on the left side. Without ACC filtering, noisy frames are still included in the offered selection set (cf. third row).

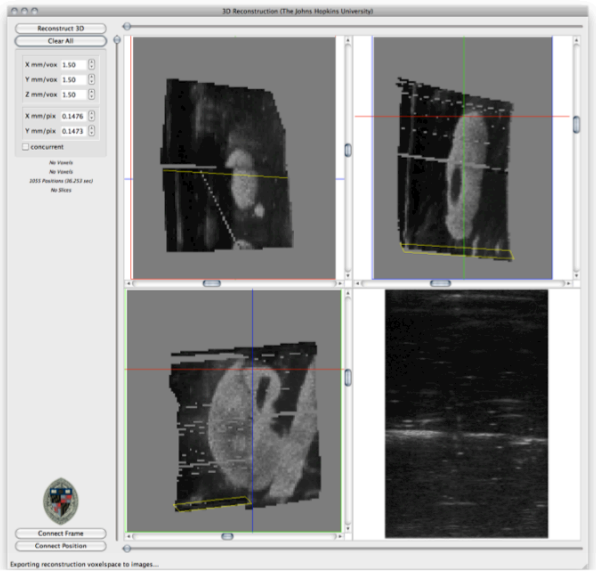

Figure 7: 2D orthogonal MPR views of a concurrently reconstructed 3D-B-mode volume (*3DR/B-Mode*) of a synthetic kidney phantom with internal structures.

Finally, a third measure to increase the EI stream quality is manual frame selection ("Frame Chooser", Figure 6). The above-mentioned pre-selection steps lessen the negative impact of this otherwise tedious task. The user is presented with a collected burst of images from which he chooses "appealing" ones. Those selected frames are then asynchronously sent out for further processing, e.g. to the 3D ultrasound reconstruction stage.

2.4 3D US Reconstruction

In this step, the two disparate real-time data streams of EI frames and 6-DoF poses are recombined to allow precise 2D frame insertion into a 3D reconstruction volume. The "*3D Reconstruction/3DR Module*" connects to one source each of those two streams to accomplish this. Given the temporal uncertainties of this distributed real-time system, proper recombination has to be based on time stamps and pose interpolation. To this end, the reconstruction module caches a 60-second sequence of pose/time-stamp pairs from which it draws to linearly/slerp-interpolate a 6-DoF pose for each new incoming image frame. In particular, this approach allows decoupling of the two streams e.g. by feeding the image stream through the above-mentioned frame selection step.

The digitally acquired 2D images (from either the B-mode or EI modules, both computed directly from the US system's real-time RF data) are projected into voxels of a 3D reconstruction volume according to a pixel-nearest-neighbor (PNN) strategy with bin averaging [6]. The single slices are inserted incrementally, i.e. the reconstruction volume is expanded on-demand when pose/slice combinations fall outside the current volume. As this amounts to a single-pass real-time approach, there is no procedure implemented for concurrent or *ex-post* hole filling. This approach has amortized linear complexity in the pixel size of the input slices, while changes of the reconstruction voxel size have no discernible performance effect. Concurrent visualization is provided by 2D orthogonal MPR in the principal planes (Figure 7).

After this step, an intra-operative dataset is available in tracking coordinates and can be used to pull in the pre-operative plan in the next registration step.

2.5 3D US-CT Registration

The registration of tracked clinical 3D ultrasound volumes and CT datasets of kidneys allows finding a transformation from CT to reference EM coordinates. To test the quality of our registration approach and to allow comparison of the properties of EI and B-Mode ultrasound volumes in a more standardized way, the phantoms described above are used for registration.

2.5.1 Metric

One of the most important characteristics of a registration approach is the metric or cost function used to compute the similarity of overlapping regions in the two volumes and therefore the quality of registration. We chose Mattes mutual information (Equation 1) as our metric to overcome the problems that might arise when registering different modalities. This was necessary as there are differences in tissue representation between B-Mode/EI and CT image data, preventing the use of simpler metrics, which work on the image intensity values directly (e.g. mean squares computation).

It is based on histogram generation for both moving and fixed image. These histograms are built based on sampled voxel values and approximate the joint and marginal discrete probability distributions. We initialized the metric computation to choose 10,000 samples that were sorted into 24 bins. As the moving image is transformed during the registration process, sample values might be taken from non-grid points of the volume. Therefore, linear interpolation was used to compute the corresponding sample values. Two representative plots of the metric can be seen in Figure 8.

$$S(\mu) = -\sum_{\iota \in L_T} \sum_{\kappa \in L_R} p(\iota,\kappa;\mu) \log \frac{p(\iota,\kappa;\mu)}{p_T(\iota;\mu)\, p_R(\kappa)}$$

Equation 1: Similarity metric according to Mattes Mutual Information (MMI) comparing two images with histograms L_T, L_R, with one image experiencing displacement μ, and joint and marginal probability distributions p, p_T, and p_R.

2.5.2 Optimizer

The optimizer evaluates the similarity metric while searching the parameter space of the registration. Given the rough cost hypersurface formed by the metric (Figure 8), direct gradient descent procedures could only guarantee to find the global optimum if the registration were provided a very good initial transformation. Furthermore the global optimum is entrenched in a steep valley, which would necessitate quickly changing the step size. Thus, methods like stochastic optimization (e.g. simulated annealing), genetic algorithms, or exhaustive search should be used to search the parameter space. A point-based registration allows finding a reproducible initial registration by manually choosing three landmarks in each ultrasound and CT dataset.

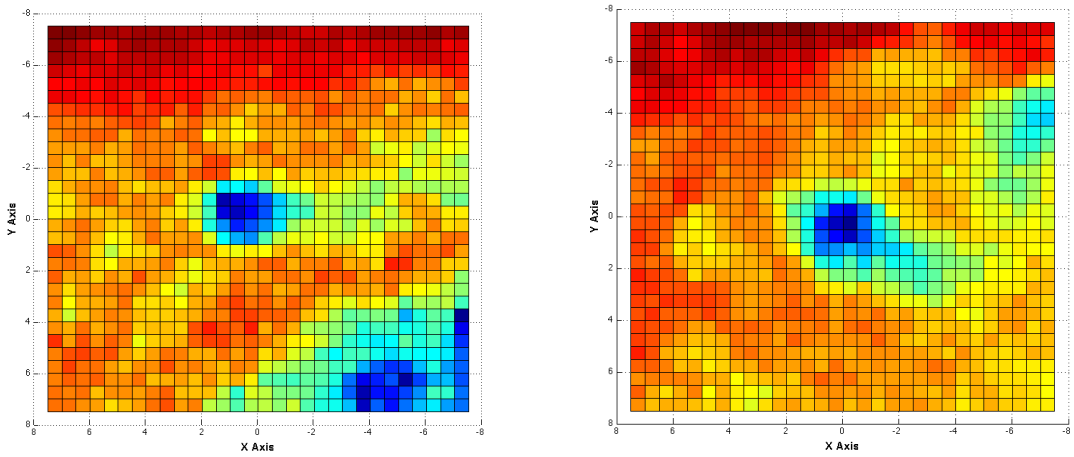

Figure 8: MMI cost function values for 3D-B-Mode/CT (left), 3D-EI/CT (right) registration under translational displacement (15 steps of 0.5mm each; function minima projected onto XY plane). Note the smoother function values for EI, with one clearly defined optimum.

As a first attempt, the presented system uses a Nelder-Meade downhill simplex optimization algorithm ("Amoeba" in ITK). Like other global optimization algorithms, this optimizer can be trapped in local optima or fail due to early convergence of some transformation parameters. Therefore it is usually necessary to adapt the parameterization of the optimizer itself to the given registration problem. The Amoeba parameters that can be modified are the initial size of the simplex around the origin (simplex delta) and the convergence tolerances for the parameters under investigation as well as for the metric values being evaluated. The optimizer is then run for a predefined number of iterations or until convergence, whichever comes first. The simplex size was initialized to values between 2 and 6, the parameter convergence tolerance set to 0.01, function convergence tolerance to 0.001, and maximum iterations to 200.

We have extended our registration approach by using the VersorRigid3DTransform optimizer for ITK, which is specialized in solving the optimization problem given by 6-DoF registrations that are described using the non-linear versor space for rotation.

2.5.3 Transformation

The search space of the optimizer for a given registration problem is characterized by the admissible transformations. The first attempt at solving the registration problem focused on translation only [Keil09] and is now extended to a six-DoF parameter space. With the assumption that only minimal deformation occurs, only rigid transformations are applied. Nevertheless, one has to keep in mind that possible deformations due to strain application for EI computation cannot be compensated only by rigid translation transformations.

After this registration step, the preparatory steps are completed and data can be passed to the modules running during the actual intervention itself (for more details cf. [4]).

2.6 CT Segmentation and Tumor Model Generation

Manual pre-operative segmentation of the CT data is performed to define the contour and size of the tumor that is to be resected, using two different techniques. Using the first method, the user paints the tumor into the slice using a circular brush with adjustable size. The second approach enables the user to outline the tumor by outlining it with the pointer and automatically defining the tumor as the area inside the shape. Both techniques are applied on the 2D slice views of the volume. The user does not have to define the tumor in all of the slices, as a trilinear interpolation determines the tumor shape on intermediate slices.

After segmentation of the tumor a surface mesh model is generated using the Marching Cubes algorithm. As the resection or safety margin is an important guidance feature for the surgeon, a further step automatically scales the tumor model by a user-defined distance to create the surface model. Both models can then be stored and used in the navigation part of the system.

2.7 Intra-operative Tracking and Visualization

To provide a live "augmented reality" view of the laparoscopic intervention with an overlay of the tumor region, it is necessary to track the camera and the tumor itself. For the latter, an EM sensor is inserted into the kidney next to the tumor capsule and fixed using an anchor mechanism attached to the sensor itself using biocompatible glue (Figure 9). It is important not to puncture the tumor capsule itself; this can be achieved by monitoring the progress of the interstitial needle containing the sensor-anchor combination with both handheld B-Mode and EI ultrasound guidance.

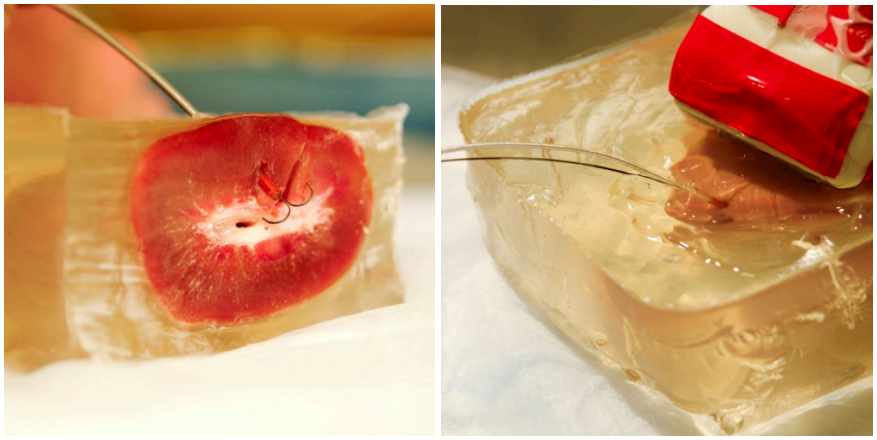

Figure 9: Sensor-anchor (left) insertion into a porcine kidney phantom under handheld ultrasound guidance (right). The anchor fixates the sensor in the tissue.

The digital video feed and all position and orientation data of each sensor are collected and sent to a central workstation running the real-time visualization module based on the medical imaging interaction toolkit (*MITK*).

The tumor model as well as a safety margin that were previously segmented from the pre-operative CT are then loaded into the MITK-based module. Each model's position and orientation in tracking coordinates, known from the registration process described before, are stored in a transformation matrix. This matrix describes the basic transformation to the models before a tracking rotation or translation can be added. To generate a 3D rendering scene of the models, a 3D render window is created by the visualization toolkit (*VTK*). With each movement the sensors attached to camera and tumor update the relative transformations that are applied to the render models. An image that shows the transformed models is now rendered and superimposed to the live video feed, thereby creating an augmented reality visualization (Figure 10).

Apart from the 3D tumor model visualization the system also provides an overlay of two concentric circles at the tumor position. These circles describe the size and position of both the tumor and the resection margin. The diameter of the circles is changed to match the apparent size of tumor and resection margin based on their distance to the camera.

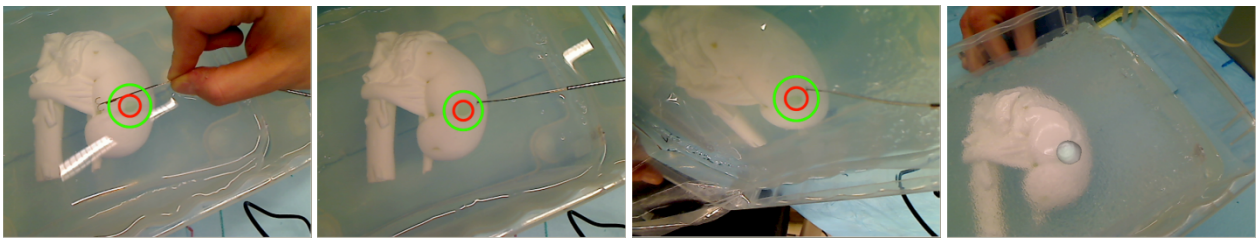

Figure 10: Real-time "augmented reality" overlay of tracked target region onto intra-operative video stream from a tracked webcam. An EM sensor (left) was manually inserted under US guidance into the tumor ROI after 3D US acquisition, with its position projected back into camera coordinates (left center and right center). The same position can be used to overlay a 3D surface model of the segmented tumor region (right).

2.8 System Architecture

The system follows a modular, distributed architecture. By breaking down the solution into a collection of independent, communicating modules, we achieve efficient resource allocation across different computers as well as decoupling of modules to minimize the impact of processing performance discrepancies. In its current implementation and

configuration, the modules span up a directed, acyclic communications graph as nodes, with communication channels like network and shared memory (depending on distribution vs. co-location and one-to-one vs. one-to-many configuration) as edges.

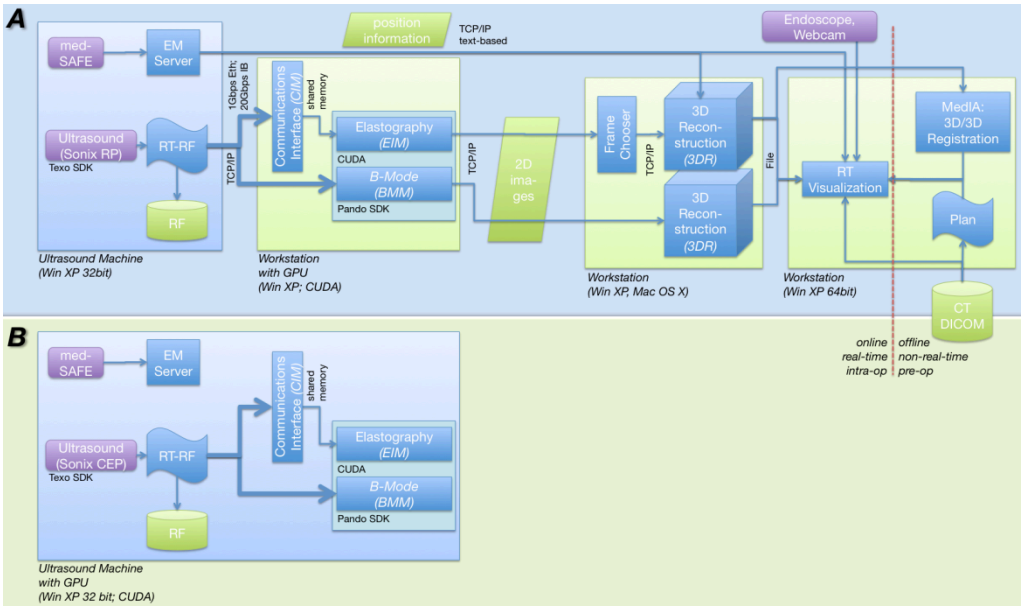

Figure 11: Proposed system software architecture, implemented in two different system setups. In the more comprehensive Setup *A*, 2D ultrasound data (left) is combined with position data into a 3D reconstruction (center), which can then be registered with CT (right) and visualized. Setup *B* represents a separate, standalone 2D B-Mode/EI workstation.

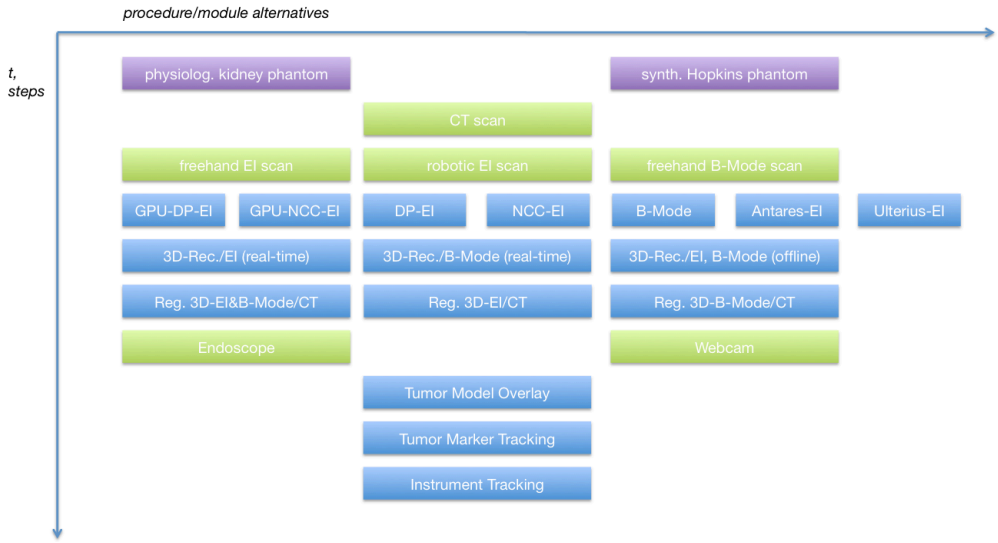

Figure 12: Process steps (vertical) and procedure/module alternatives (horizontal). All these combinations are possible with the proposed architecture.

While a few constraints were given for the current system configuration (due to ultrasound acquisition hardware and various hardware drivers tied to certain operating systems), most of the remaining features had no such restrictions. Therefore, a functional decomposition of the system requirements led to the module set and architecture as shown in Figure 11. An obvious benefit is the ability to distribute the computation (e.g. for local vs. remote visualization or remote strain computation) according to the current application's needs. Two such implementations are shown in the figure:

Setup A includes the components necessary for a complete "tracked freehand 3D ultrasound for CT-based navigation" system as described in the previous paragraphs. This system can be easily distributed across different machines to accommodate limited space in the operating room (*OR*), higher computation requirements for online strain computation, and the need for more screen real estate during the offline planning phase. As a second system, *Setup B* uses a subset of those modules to provide a stand-alone "2D B-mode/strain imaging" workstation. Tracking data may optionally be collected for later 3D reconstructions, if required. A variety of possible procedures and underlying module combinations can be seen in Figure 12.

The architecture is very flexible in the sense that any ultrasound machine can be the imaging source, as long as it can export RF data (either in real-time or stored to disk), ideally with associated timestamps. Similarly, different tracking mechanisms can be employed, such as EM tracking like in the presented system, but also optical tracking or systems based on local sensors such as opto-inertial components as in the *BeeSpaceMouse* [7]. Strain image computation can be achieve through a multitude of different modules – NCC, DP, GPU-based variations of both, but also with the respective proprietary modules available in commercial (Siemens, Ultrasonix) systems. Intra-operative 3D-US/CT registration can be performed using either B-mode or EI. Most of the components not tied to specific hardware have been ported between different operating systems (Windows XP 32-bit, 64-bit, Linux (kernel 2.6), and Mac OS X 10.6, as well as several different CUDA GPU devices).

Finally, we also plan to offer the infrastructure framework as well as particular module implementations available for download from our website in the near future.

3. RESULTS

3.1 Experimental Results

A variety of experiment series (with tracked US in breast irradiation planning [2], in ablation monitoring [5], and robot-based US palpation motion optimization [4]) has each shown an increase in elastography imaging quality. Combined, they promise robust handheld laparoscopic EI acquisition.

EI slice acquisition can be performed either using a 3-DoF robot-based setup (to ensure repeatable palpation motion parameters and micrometer-precise slice position information) or in an EM-tracked handheld approach. With the robotic setup, EI images were acquired at maximum decompression of a palpation cycle to consistently minimize tissue deformations. For ultrasound acquisition, both Ultrasonix Sonix RP/CEP (for B-Mode and RF data) and Siemens Antares (for B-Mode and EI data) machines were used. By parallelizing both (NCC and DP) algorithms using CUDA, it was possible to increase the frame throughput by a factor of 6…10 as compared to the corresponding C implementations. While this amounts to approx. 300fps for NCC-EI and 30fps for DP-EI, these numbers describe "batch-mode" usage, where all RF frames are pre-loaded into GPU memory [1]. For real-time usage where RF data and EI frames are continually moved between host and device memories, a continuous frame rate of approx. 5…10fps was achievable for a non-optimized input/output setup.

The **3D volume reconstruction** operates at a speed of ~24 slices/*sec* (with a single thread on a 2.66GHz Intel Core 2 Duo with 6MB L2 cache), independent of the type of the input images, i.e. B-mode or EI. The reconstruction voxel size was set to a range of values between 1.0*mm* and 2.0*mm*, with little discernible effect on achievable frame rate or on the subsequent process steps.

The final volume is transferred to the **rigid registration** algorithm to align with CT data (using Mattes Mutual Information/MMI as the similarity metric). The **registration** application is based on the open source development toolkits MITK from the German Cancer Research Center (DKFZ), ITK and VTK from Kitware Inc., as well as Qt from Nokia/Qt Software. Using these toolkits enables working with standardized, well-established algorithms. The registration experiments were performed using subvolumes cropped from the EI, B-Mode, and CT volumes. The region of interest was chosen based on the expected amount of data for our clinical application. As only a small portion of the kidney is scanned during the operation, the experimental volumes were comprised almost exclusively of kidney tissue. In order to facilitate a comparison between the registration properties of strain images and B-Mode images, both datasets were registered with CT using different initial offsets for the registration – single displacements along the *x* and *y* axes as well as all translational perturbations by +/–2*mm* and +/–4*mm* in all three dimensions. The last test was perturbation by +/–2*mm* along the *x* and *y* axes and 4*mm* in *z* direction, respectively. Given the general restrictions in the use of ex-vivo

tissue for elasticity imaging, the geometry and availability of features in the physiological porcine kidney phantoms approximated in-vivo conditions.

Registration of the volume of interest takes 5–10 seconds for three translational degrees of freedom (extension of registration to six DoF is discussed in an upcoming publication). While the cost function values of MMI between EI and CT show a relatively smooth behavior (Figure 8), the capture range is nevertheless fairly small, making pre-registration necessary.

The quality of the **registration results** was measured using visual examination using the fiducials that were added to the phantom. Those fiducials were clearly visible in strain, CT and B-Mode images. For getting correct registration results for the B-Mode case, the registration parameters often had to be optimized very specifically to the given problem. As there are several local minima having equal of even smaller (i.e., better) values for the metric, it is hard for the Amoeba optimizer to find the minimum of interest. The situation is much different for the strain to CT image registration, where the cost function has only one very pronounced global optimum. However, this global optimum can also be missed by the Amoeba optimizer under certain conditions (e.g. initialization of the simplex fully on the plateau).

In order to account for the different imaging modalities, preliminary experiments with other metrics were performed as well. In particular, noting that B-Mode US shows gradients of acoustic impedance, the gradient of the CT data was computed to match this property. Then, both images were registered using a mean-squares error metric. However, this approach proved less effective and produced even larger registration errors than the MI-based approach discussed here, in spite of the tissue representation being more similar in the two modalities then.

These results show that the metric used for the chosen registration algorithm is very well-suited for estimating the similarity between CT and EI images, even for the small field of view given in our test data.

The medSAFE **tracking system** consists of a 9-axis flat magnetic field transmitter and an electronics unit, supporting up to four wired 6-DoF sensors simultaneously. Each sensor is connected to the electronics unit via a pre-amplifier. The system can acquire up to 200 position measurements per second within a measurement volume of $400mm \times 400mm \times 360mm$. The sensors attached to camera and tumor are of $1.3mm$ diameter and $7.7mm$ length. The tracking system can achieve a RMS error $<1mm$ when moving the sensor at $<200mm$ per second. The cable connecting the sensor to the electronics unit of the tracking system has a diameter of $1.2mm$ and measures $2m$ in length, making the cable thin enough to fit through a trocar while being attached to an endoscopic camera.

Currently a Logitech QuickCam camera provides 640×480-pixel **live video** footage. In future system designs the video source will be replaced by a laparoscope and therefore will meet the requirements for minimally invasive surgery. For a smooth and real-time video navigation experience we are calculating at least 25 augmented video images per second. The much higher rate of position information from the tracking system enables outlier correction, which will be implemented in future system stages. Similar **tracking/visualization** system designs using both optical and magnetic tracking systems are proposed in literature, e.g. in [3]. Nevertheless there are key differences; the most important being that the video overlay in the presented system can be done not only in real-time, but also online and not retrospectively.

3.2 Discussion

With the proposed system, it will be possible to perform safe partial nephrectomies by leveraging pre-operative planning data. The ability to reliably avoid blood and collective vessels while honoring safe resection boundaries, with only a small additional hardware footprint, promises good acceptance of the system in clinical environments.

While the main drawback versus the current standard laparoscopic procedure is the manual acquisition of the 3D ultrasound volume, this constitutes a minor time investment in practice. The principal benefit of the system has to be seen in the “heads-up display” capability of overlaying the tumor position onto the live endoscopic video stream, instructing the surgeon as to where cutting is safe. In particular, this achieves a vast reduction of preparation time while clamping the renal artery and interrupting blood flow, because pre-operative planning can be executed directly in the OR, allowing the surgeon to concentrate on a quick resection of the tumor mass.

3.3 Outlook

The system shows that the main problems – robust real-time elastography and robust 3D-US/CT registration – are solvable. What needs to be evaluated is how well the phantom and laboratory results transfer into the animal testing and clinical stages. To achieve this, accuracy, reliability, and usability will need to be investigated.

With the software module set's rapid expansion in our group, the software architecture is seeing equally strong development. After appropriate preparation of the implementations, we expect to make the discussed modules available for public in the near future.

3.4 Acknowledgements

Further support was provided by internal funds of the Johns Hopkins University and by the Fraunhofer Gesellschaft within the PROFIL program. Equipment support was generously provided by Ultrasonix (Sonix RP/CEP ultrasound), Siemens Medical Systems (Antares ultrasound), Dymax Europe (Medi Cure biocompatible glue), and the ERC-CISST (robotic stage). We also gratefully acknowledge the research efforts and extensive development contributions of our students, in particular of Hyun-Jae Kang and Nishikant Deshmukh.

REFERENCES

[1] Deshmukh N et al.: *"GPU-Based Elasticity Imaging Algorithms"*, MICCAI Grid Workshop (2009).
[2] Foroughi P et al.: "*Multi-Modality Fusion of CT, 3D Ultrasound, and Tracked Strain Images for Breast Irradiation Planning*", Proc. SPIE Med Img (2009).
[3] Nakamoto M et al.: "*Realtime Organ Tracking for Endoscopic Augmented Reality Visualization Using Miniature Wireless Magnetic Tracker*", Proc. MIAR'08 – 4th International Workshop on Medical Imaging and Augmented Reality (2008).
[4] Keil M et al.: "*Ultrasound and CT Registration Quality: Elastography vs. Classical B-Mode*", Proc. ISBI (2009).
[5] Rivaz H et al.: "*Ablation Monitoring with Elastography: 2D In-Vivo and 3D Ex-Vivo Studies*", Proc. MICCAI (2008).
[6] Solberg OV et al.: *"Freehand 3D Ultrasound Reconstruction Algorithms – A Review"*, Ultrasound in Med. & Biol. (2007).
[7] Stolka Ph et al.: "*Multi-DoF Probe Trajectory Reconstruction with Local Sensors for 2D-to-3D Ultrasound*", Proc. ISBI (2010, accepted).