Computational methods for minimally invasive image-guided interventions

PhD thesis summary
András Lassó

Advisor: Dr. László Vajta
2011
Introduction

For centuries physicians used their senses directly, mainly sight and touch to diagnose and treat illnesses. The emergence of X-ray radiography in the first decades of the 20th century, followed by several new medical imaging modalities, such as X-ray computed tomography, ultrasound, magnetic resonance imaging enabled the physicians to acquire accurate structural information from the human body without direct physical interaction. Development of imaging technology, coupled with the evolution of medical instruments, soon led to the appearance of a new class of medical procedures: minimally invasive image-guided interventions. In these procedures both sensing and intervention are performed indirectly, through the use of various devices and tools in a minimally invasive manner, i.e., causing minimum damage to the healthy tissues.

The main challenges in image-guided intervention research are high-resolution visualization and navigation, efficient image acquisition, registration, and segmentation, handling intraoperative tissue deformations, and standard methods for performance validation ([Haigron2009], [Cleary2010], and [DiMaio2007b]).

Research objectives

Having studied telerobotics and working on image-guided interventions for many years, I focused my attention on challenges related to generic modeling of these interventional systems and developing specific image enhancement and registration methods.

I had the opportunity to participate in the development of the GE Innova® interventional X-ray system. My goal was to create a new method that can be used in these systems for improving the visibility of blood vessels without additional X-ray or contrast agent dose, or achieving equivalent image quality with lower dose, by utilizing the temporal information contents of the acquired image sequences.

Later, I also participated in creating a robot-assisted magnetic resonance imaging (MRI) guided prostate biopsy system. This system is used for extracting tissue samples from the prostate for cancer detection. Dislocation of the prostate was found to be the major source of biopsy targeting error. The objective of my work was to understand and quantify the prostate motion, and to create a compensation method that enables good targeting accuracy even if the subject moves.

While developing and working with different image-guided intervention systems I realized that although some low-level hardware and software components are shared between different implementations, and design requirements of these systems were very similar, each system was designed from the ground up. This could be explained by that no accurate generic system model was available for these systems. Without a common model sharing and reuse between different systems were very difficult. My objective was to create a system model for image-guided interventions that can serve as a common architectural basis for various system implementations.
Contributions of the thesis

Thesis group 1: Vessel enhancement by temporal statistical learning

Motivation

Morphological or anatomical information about the vascular system is essential for detection of vascular diseases and planning of surgical procedures or catheter interventions. Vessels, organs of interest and tools have to be displayed with the highest spatial and temporal resolution and fidelity so that their size, relative position and temporal changes can be estimated accurately, while irrelevant background structures (bones, muscles, etc) should be suppressed.

In X-ray angiography (XA) a radio-opaque contrast agent is usually injected into the vessels of interest by means of a catheter. The contrast agent travels through the vessels, making them visible in the X-ray, and is eventually washed out by the blood stream. An X-ray sequence capturing the passage of the agent through a given vessel region gives information about the vessel morphology and the dynamics of the blood flow.

Most commonly used methods, such as the Digital Subtraction Angiography (DSA) tend to utilize only a fraction of the available spatiotemporal information when computing morphological information from an XA image sequence. The use of temporal information (i.e., the evolution of the grey level values in time) is generally limited to simple averaging and subtraction operations on selected frames.

Temporal bandpass filtering (TBF) methods have been proposed as an alternative to DSA including temporal dynamics ([Kruger1979], [Kruger1981], [Kruger1982], [Riederer1983], [Liu1985], and [Kump2001]), with the aim of achieving high-quality images with better dose efficiency. TBF methods model the temporal dynamics through recursive filters ([Kruger1981]) or matched filters ([Kruger1982]), using filters designed manually from observations in example sequences.

However, the crucial problem for temporal methods is how to model the evolution of grey levels with sufficient generality. Kump et al. [Kump2001] used approximate matched filtering (AMF) consisting of a correlation with a kernel that approximates an optimal TBF, followed by a maximum opacity operation. They reported that the AMF nearly achieved the performance of the optimal TBF on simulated images. However, the maximum opacity operation makes this method sensitive to noise and motion present in clinical images, and user interaction is needed to mark vessel points for obtaining filter kernel estimates.
Contributions

The temporal evolution of grey levels is hard to capture by a hand-crafted model, but it can be learnt automatically from examples.

1.1 I developed a new method—support vector machine temporal filtering (STF)—to enhance vessel visibility in X-ray angiography sequences using statistical learning methods.

The proposed method relies on the usage of support vector machines (SVM), which have been successfully applied to a wide range of classification and regression problems thanks to their good generalization performance, efficiency and robustness ([Vapnik1998]). SVM is used first to learn the temporal evolution of gray levels at vessel pixels from training sequences, then to classify pixels as “vessel” or “not vessel” in the same sequence, or in other sequences that were acquired with the same examination protocol. I showed that the distance of a sample from the hyperplane that optimally separates sample vectors is equivalent to the output of an optimal matched filter for the XA image, in the sense that the contrast-to-noise ratio is maximized. Therefore, an SVM with a simple linear kernel is applicable to distinguish vessel and non-vessel pixels from the temporal evolution of their grey levels, and the signed distance from the optimal separating hyperplane is a meaningful metric for “vesselness” (probability of the point being in blood vessel).

The dynamics of the propagation of the contrast agent in the blood flow can differ significantly, but it is noticeable that the factors that influence the blood flow (such as vessel diameter, blood velocity) change smoothly in space. This suggests introducing a spatial localization constraint, whereby only spatially close training pixels play a role in the classification of a sample. This leads to improved visibility, as the classification is not disturbed by distant, potentially different gray level dynamics. The locality constraint can be naturally built into the SVM framework by constructing an appropriate non-linear kernel.

1.2 I created a new kernel function to incorporate a locality constraint into the SVM framework, which improves vessel visibility simultaneously in both low-contrast and high-contrast image regions:

\[
K_{RL}(\mathbf{x}, \mathbf{z}) = \exp\left(-\gamma \sqrt{(x_1 - z_1)^2 + (x_2 - z_2)^2}\right) \sum_{i=1}^{N} x_{i+2} z_{i+2},
\]

where the first two elements of the sample vectors \((x_1, x_2, z_1, z_2)\) are the spatial coordinates, and subsequent elements \((x_3, x_4, \ldots, z_3, z_4, \ldots)\) the grey levels. The parameter \(\gamma\) controls the localization strength: the smaller its value, the more distant training vectors influence a specific point. For \(\gamma = 0\) the proposed new kernel becomes equivalent to a simple linear kernel.
**Application of the results**

The vessel visibility enhancement method was extensively tested on synthetic data and also on various clinical image sequences from different parts of the body, containing different types of motion and blood-flow dynamics.

The performance of the proposed STF method was compared to five different vessel enhancement methods evaluated in [Kump2001] by computing the contrast-to-noise ratio (CNR), which characterizes the vessel visibility. The proposed STF filter performed slightly better than the best already existing method (AMF), achieving 2 to 12% higher CNR. Moreover, the AMF method is sensitive to the presence of structured noise (such as motion artifacts), while STF suppresses most of them. Tests showed that the presence of structured noise did not impact the STF processing results, while CNR dropped by about 25% when the AMF method was used.

The STF method was evaluated on multiple clinical images. Figure 1 shows an example processing result on a clinical image with significant motion. While processing with subtraction resulted lots of motion artifacts and low vessel contrast, the STF method efficiently suppressed the moving background and clearly visualized the blood vessel.

![Figure 1. Comparison of improved vessel visibility on a clinical image using subtraction (left) and STF (right). STF suppressed the moving soft tissues significantly better than the subtraction method.](image)

The vessel enhancement method has numerous potential applications, such as segmentation of coronary vessels for cardiac blush assessment [Ungi2009a], [Ungi2009b], and [Ungi2009c] or brain vasculature for fluid-structure-interaction analysis [Lasso2008a]. A variant of the method was also evaluated on tissue classification in contrast enhanced digital mammography images ([Jeunehomme2004]) and ultrasound-guided tissue ablation monitoring ([Imani2011]).

**Publications related to the topic of this thesis group**

- [Lasso2005]: Journal paper describing the STF method and its application for XA image sequences.
- Other papers using SVM for medical image segmentation:
  - [Jeunehomme2004]: Conference paper, describes application of SVM temporal filtering with a simple linear kernel for contrast-enhanced mammography sequences.
  - [Imani2011]: Journal paper describing a method that uses SVM for determining ablated tissue regions using ultrasound RF time series images.
• Other papers on XA image processing:
  o [Ungi2009a], [Ungi2009b], [Ungi2009c]: Two journal papers and a conference abstract describing a quantitative blush assessment method including a vessel segmentation step. STF could have been used on these images for vessel segmentation, but it was not necessary, as acceptable results were achieved with a simpler method.
Thesis group 2: Motion characterization and compensation for MRI-guided transrectal prostate biopsy

Motivation

MRI provides a promising approach to the detection and treatment of prostate cancer. It has high spatial resolution, excellent soft tissue contrast, and volumetric imaging capabilities [Tempany2008].

Krieger et al. developed an MRI-guided transrectal robotic prostate biopsy system ([Krieger2005], the manipulator is shown in Figure 2) that has been used in over 200 biopsies to date at the U.S. National Cancer Institute (Bethesda, MD, USA). At the beginning of the procedure a high quality MRI image of the patient’s pelvis is acquired (referred to as the target planning volume) and the physician determines the biopsy target locations on this image. The following three steps are then performed for each planned biopsy targets:

1. The robot aligns the needle towards the target position and the physician inserts the needle.
2. After the needle is in place, a quick, lower-quality MRI image is acquired for confirmation (referred to as post-needle insertion volume). This step may be omitted (depending on the available time, suspicion that the patient moved, etc.).
3. The biopsy core is extracted and the needle is retracted.

Figure 2. Side view of the MRI-compatible intra-prostatic needle placement device developed by Krieger et al. (Source: [Krieger2005])

Often, the post-needle insertion image shows that the actually reached position visible does not match the originally planned biopsy location. This error is mainly caused by prostate motion and deformation ([Krieger2005]).

To improve targeting accuracy during these procedures, it is necessary to characterize the patient motion and prostate deformation and develop a computational method to compensate the effect of the deformations.
Contributions

2.1 I developed a three-stage deformable image registration method for quantitative characterization of prostate motion during robot-assisted MRI-guided transrectal biopsy procedures.

The registration method uses the target planning and post-needle insertion volumes as inputs. The method is composed of a pre-processing step and three registration stages. An overview of the method is shown in Figure 3. In the first registration stage the target planning volume was aligned to a region of interest (ROI) in post-insertion volume by a rigid transformation. This stage compensated for prostate motion in coherence with the device and patient. To correct for residual decoupled prostate motion, the resulting image was registered again in the second stage with the original fixed image using only the prostate as the region of interest. The third stage of the registration used a non-rigid transformation, defined by equally spaced control points in the prostate volume. The deformation at any point was computed by interpolating between the dislocation vectors defined at the control points using a B-spline kernel.

The proposed method is built on some well-established registration schemes and algorithms and contains the following new concepts:

1. Appropriate ROI was defined for each registration step (identified constraints and specified a generally usable solution);
2. Appropriate similarity metric was determined (chose a suitable metric and modified it for making usable for small ROIs);
3. Suitable deformable transform model was specified (chose appropriate transform type and determined usable transform parameters).

Volume-to-volume registration methods cannot be applied to determine prostate motion during an interventional procedure, as acquisitions time for volumetric images is prohibitively long. However, it is feasible to acquire a few image slices during various time of a procedure, as such an acquisition takes only a few seconds to complete. A method that can recover the prostate motion from only a few image slices may fit into the interventional workflow and be utilized during interventions.

2.2 I created a new method for improving targeting accuracy during robot-assisted MRI-guided prostate biopsy procedures by motion compensation. The method uses two-stage multi-slice-to-volume registration to align intra-procedural slice images to the target planning volume to recover prostate motion.

The registration method is composed of a pre-processing step and two registration stages. An overview of the method is shown in Figure 4. A rigid registration was first performed to obtain an initial pose of the target planning volume, which was then non-rigidly registered to the fixed sparse volume. A sparse image was constructed from the available image slices, which allowed the use of methods that were originally developed for volume-to-volume registration. Similarity metric, deformation model, and optimizer algorithms were adopted from the volumetric registration method.

This slice-to-volume registration method has the following unique combination (compared to other similar algorithms described in the literature) of features:

1. It uses multiple image slices for the registration instead of a full volumetric image,
2. It supports non-rigid transformation, and
3. It utilizes the prostate motion characterization results (for determining transformation parameters and region of interests, and for validation of the method).

Applications of the results

The prostate motion characterization and compensation methods were tested on images with simulated deformations and on real clinical images. Qualitative and quantitative validation of the volume-to-volume registration method indicated a maximum registration error of about 2 mm (mean Hausdorff distance between the registered organ surfaces). Motion characterization was successfully performed on clinical images of more than 80 biopsies. Displacement of biopsy targets inside of the prostate was computed.

The intra-procedural prostate motion compensation method was tested on 30 simulated and multiple clinical images using three orthogonal intra-procedural slices. In simulated images the difference of the known ground truth and the computed dislocation (the target registration error, TRE) was computed for all target positions in the prostate. The results showed that only using three slices the registration error in the prostate region was almost as good as when using full volume-to-volume registration. The slice-to-volume registration reduced the initial (without registration) TRE of 2.1–5.6 mm to as low as 0.6–0.9 mm. Evaluation on real patient data sets showed that the method achieved both the accuracy (<3 mm error was required; about 1–2 mm was achieved) and computation time limit (<60 seconds was required; about 40s was achieved) that are required for interventional use. A sample registration result is shown in Figure 5.

![Figure 5. Prostate, rectum, pubic bone contours overlaid on the target planning volume: without registration (left), after the rigid stage of the slice-to-volume registration (middle), and at the end of the slice-to-volume registration, including the deformation stage (right).](image-url)
Publications related to the topic of this thesis group

- Prostate motion characterization (Andras Lasso’s main contributions were the development and validation of the registration method; Helen Xu’s main contribution was the application of the registration method for retrieving and analyzing the prostate, needle, and patient motion)
  - [Xu2010b], [Xu2010c], [Xu2010d]: Conference paper and two abstracts on prostate characterization (2-stage registration, rigid only)
  - [Xu2010a]: Journal paper on prostate characterization (3-stage registration, deformable)
  - [Lasso2010b]: Conference paper describing the FEA-based registration validation framework.

- Prostate motion compensation using multi-slice-to-volume registration (Hadi Tadayyon developed initial versions of the motion compensation method with help from Andras Lasso, Andras Lasso’s main contributions were 1) the complete rework of the method to make it much simpler and much more efficient, 2) validation of the method):
  - [Tadayyon2010a]: Conference paper on an early version of the prostate motion compensation method (uses 1-stage rigid registration only, tested on non-deformed simulated and phantom images)
  - [Tadayyon2011] (submitted), [Tadayyon2010b]: Journal and conference paper on an improved version of the motion compensation method (uses 2-stage, rigid and non-rigid registration). The method described in the thesis is an optimized (simpler, much faster) version of the method described in the papers.
  - [Lasso2010b]: Conference paper describing the FEA-based registration validation framework.

- Papers related to the MRI-guided prostate biopsy system where the motion compensation method is applicable:
  - [Lasso2009], [Lasso2011a], [Lasso2011c]: Conference papers focusing on the software system architecture and implementation.
  - [Tokuda2011b] (submitted), [Tokuda2011], [Tuncali2011]: Journal paper and two conference abstracts, focusing on clinical application and results.
Thesis group 3: System model for minimally invasive image-guided interventions

Motivation

Most of the devices used for minimally invasive image-guided procedures can be considered as a special class of remotely controlled manipulators, or telerobots. The telerobots are controlled by humans, but typically the operator’s high-level requests are translated to actual manipulator motion by a computing system (human-interactive computer, HIC), which takes into account the capabilities of the manipulator, predefined motion constraints, and sensor feedback. Motion constraints can be related to keeping the skin entry point position unchanged or protecting sensitive regions inside the body from interacting with the manipulator. The HIC also provides rich, integrated display of all relevant information to the operator. The communication channel between the HIC and the manipulator is typically limited (e.g., by low bandwidth, large delay, unreliable transfer, loosely integrated system components), therefore direct control of the manipulator is accomplished by another computer, the task-interactive computer (TIC). This interaction of human and computer to accomplish tasks corresponds to the model of human supervisory control (HSC), described by Sheridan ([Sheridan1992]).

The HSC model assumes that all sensors provide continuous real-time data. However, in image-guided intervention systems there are imaging and other data acquisition modalities that cannot provide continuous real-time feedback because of physical or technological limitations or due to lack of tight integration between system components. This heterogeneous system of tightly integrated and loosely coupled, real-time (RT) and non-real-time (NRT) components (heterogeneous image-guided intervention system) is very typical in the clinical environment. Typical real-time sensors are position sensors (embedded in or attached to devices or the patient) and certain imaging modalities (e.g., some ultrasound, X-ray fluoroscopy, and optical imaging systems). Examples of non-real-time data include volumetric images from ultrasound, X-ray, computed tomography (CT), and MRI (acquisition times ranging from a few seconds to several minutes). Images that can be acquired real-time but needs long time to transfer (due to software or hardware interface limitations) or process (to extract signal that is required for the functioning of the system) also have to be considered as non-real-time data.

Having a common system model allows the development of standard, optimized solutions for the most frequent and/or most difficult challenges in building a complete system. Problem analysis, design, and implementation have to be performed only once, and the results can be applied to many systems. However, the conventional model of human supervisory control gives an inaccurate representation for image-guided intervention systems, which are almost always heterogeneous in the sense that they utilize both RT signals and NRT signals.
Contributions

3.1 I extended the conventional human supervisory control model to create a heterogeneous human supervisory control (HHSC) model to yield an accurate representation of heterogeneous minimally invasive image-guided intervention systems. Applying the HHSC model enables utilization of all real-time and non-real-time information acquired during the image-guided procedure.

An overview of the model is shown in Figure 6. An additional set of NRT sensors were added to the basic HSC model. The RT and NRT signals may convey complementary, redundant, or contradictory information, thus, they have to be consolidated before being used for controlling the effector. This consolidation is a complex data fusion task, often requiring inputs from the human operator.

![Figure 6. Heterogeneous human supervisory control (HHSC) model](image)

3.2 I created a new data representation scheme that can be used to store multiversion spatiotemporal data in a scene graph using matrix nodes. The scheme allows complete representation and flexible exploration of all data acquired during an image-guided interventional procedure along all spatial and temporal dimensions.
The matrix node has 3 spatial + 2 time dimensions (Figure 7). It has two time indices: valid time and original time. When new version of a spatial object is acquired (e.g., a new image is acquired, or a modification is performed in the geometry of a model), then this new spatial object is placed into a new row of the matrix. One row of the matrix describes how that particular version of the spatial object evolves in time. One column of the matrix contains all the versions of the spatial object in the position that is valid at the corresponding valid time. If a sensor can provide continuous real-time signal then each matrix element stores a continuous stream of data, which is valid between two time points.

Figure 7. Matrix structure for storing temporally changing spatial data. The data stored at \((r, c)\) position in the matrix corresponds to the valid state at time \(c\) of the data acquired at time \(r\). The matrix structure allows review of the data along different spatial and temporal dimensions.

By creating a new row for each modification of the spatial object, all the states of the object are maintained in the model: no data is discarded or overwritten. This guarantees that all information acquired during the procedure is preserved.

The matrix structure allows navigation in the spatiotemporal data set along various dimensions (Figure 7). By showing the scene at a specific valid time while modifying original time the temporal changes of the spatial objects are visualized. This can be used to show gray level changes in an image with motion compensation, or review changes of the intervention plan during the procedure overlaid on the original planning image. Similarly, spatial changes of an object can be reviewed by fixing original time and varying valid time. Object versions can be reviewed in their original state by modifying both time indexes synchronously (keeping original time = valid time).
Application of the results

The described model was used for implementing planning and control software for a robot-assisted MRI-guided transrectal prostate biopsy system. The biopsy robot is equipped with electro-optical joint encoders on the needle guide, which enables real-time tracking of the expected needle position ([Krieger2011]) and is prepared for supporting fully automated needle insertion using actuated joints.

The planning and control software was implemented as an extension module for the 3D Slicer ([Pieper2004]) medical image analysis and visualization application framework. A sample of the user interface of the application is shown in Figure 8. The software is already in clinical use, biopsies on real patients are performed at U.S. National Cancer Institute (Bethesda, MD, USA), Johns Hopkins Hospital (Baltimore, MD, USA), and Harvard’s Brigham and Women’s Hospital (Boston, MA, USA).

![Figure 8. Screenshot of the planning and control software used for MRI-guided transrectal biopsy.](image)

The applicability of the HHSC model was demonstrated in two other interventional systems (ultrasound-guided focal brachytherapy of the prostate; X-ray-guided intracardiac robotic catheter ablation), which illustrates that the HHSC model can describe a wide range of image-guided intervention systems, using various imaging modalities, robotic devices, and target anatomies.

Publications related to the topic of this thesis group

- [Lasso2011a], [Lasso2009]: Conference papers focusing on architectural and system modeling issues for MRI-guided robot-assisted prostate biopsy.
- [Urbancsek2000]: Related journal paper describing a tele-microrobot system model, developed for biotechnology research.
Publications

Journal papers


Submitted:

MICCAI conference papers

MICCAI is the highest ranked conference in the field of Medical Image Computing and Computer Assisted Intervention. Papers are 8 pages long and peer-reviewed in a two-tier double-blinded process. MICCAI papers, published annually in a 2-volume book by Springer, carry the effective impact factor of a journal, considering the number and weight of references to MICCAI papers in other publications. MICCAI is indexed on PubMed.


Other conference papers


Conference abstracts


Patents


Other publications


Bibliography


