Architecture of a High-Performance Surgical Guidance System Based on C-Arm Cone-Beam CT: Software Platform for Technical Integration and Clinical Translation

Ali Uneri a, Sebastian Schafer b, Daniel Mirotta a, Sajendra Nithianathan b, Yoshito Otake a, Sureerat Reaungamornrat a, Jongheun Yoo a, J. Webster Stayman b, Douglas Reh c, Gary L. Gallia d, A. Jay Khanna e, Gregory Hager a, Russell H. Taylor a, Gerhard Kleinszig f, and Jeffrey H. Siewerdsen a,b

a Department of Computer Science, Johns Hopkins University, Baltimore MD
b Department of Biomedical Engineering, Johns Hopkins University, Baltimore MD
c Department of Otolaryngology, Johns Hopkins Medical Institute, Baltimore MD
d Department of Neurosurgery, Johns Hopkins Medical Institute, Baltimore MD
e Department of Orthopaedic Surgery, Johns Hopkins Medical Institute, Baltimore MD
f Siemens Healthcare, Erlangen Germany

ABSTRACT

Intraoperative imaging modalities are becoming more prevalent in recent years, and the need for integration of these modalities with surgical guidance is rising, creating new possibilities as well as challenges. In the context of such emerging technologies and new clinical applications, a software architecture for cone-beam CT (CBCT) guided surgery has been developed with emphasis on binding open-source surgical navigation libraries and integrating intraoperative CBCT with novel, application-specific registration and guidance technologies. The architecture design is focused on accelerating translation of task-specific technical development in a wide range of applications, including orthopaedic, head-and-neck, and thoracic surgeries. The surgical guidance system is interfaced with a prototype mobile C-arm for high-quality CBCT and through a modular software architecture, integration of different tools and devices consistent with surgical workflow in each of these applications is realized. Specific modules are developed according to the surgical task, such as: 3D-3D rigid or deformable registration of preoperative images, surgical planning data, and up-to-date CBCT images; 3D-2D registration of planning and image data in real-time fluoroscopy and/or digitally reconstructed radiographs (DRRs); compatibility with infrared, electromagnetic, and video-based trackers used individually or in hybrid arrangements; augmented overlay of image and planning data in endoscopic or in-room video; real-time “virtual fluoroscopy” computed from GPU-accelerated DRRs; and multi-modality image display. The platform aims to minimize offline data processing by exposing quantitative tools that analyze and communicate factors of geometric precision. The system was translated to preclinical phantom and cadaver studies for assessment of fiducial (FRE) and target registration error (TRE) showing sub-mm accuracy in targeting and video overlay with intraoperative CBCT. The work culminates in the development of a CBCT guidance system (reported here for the first time) that leverages the technical developments in C-arm CBCT and associated technologies for realizing a high-performance system for translation to clinical studies.

Keywords: image guided surgery, cone-beam CT, intraoperative imaging, surgical navigation, augmented reality, image registration, digitally reconstructed radiograph

1. INTRODUCTION

A prototype mobile C-arm capable of high-performance cone-beam CT (CBCT) has opened new opportunities for high-precision surgery and motivated the development of new navigation tools for use with intraoperative imaging. While conventional methods rely on fixed preoperative images of the patient anatomy, the intraoperative imaging paradigm and the tools it enables make use of a continually updated image context that properly reflects anatomical change. The prototype mobile CBCT C-arm demonstrates sub-mm spatial resolution and soft-tissue visibility at low radiation dose, thus offering to advance the use of surgical navigation across a broader spectrum of complex surgeries, specifically those involving soft tissue in which preoperative data alone is insufficient for guidance in the context of tissue deformation and excision.
From a system and workflow perspective, the intraoperative mobile C-Arm providing CBCT upon request introduces an additional dynamic to system integration, specifically considering its central role as the subject of reference for other guidance techniques, e.g., surgical trackers. Integrating these systems in a flexible and dynamic manner is therefore an important, challenging task, and one that is not fully addressed by commercially available image-guidance technologies, such as StealthStation (Medtronic, Minneapolis MN) or the Kolibri and VectorVision (BrainLab, Feldkirchen Germany) systems, which have been conceived around preoperative imaging. Each leaves considerable room for improvement in terms of workflow, integration, and applicability to a broad spectrum of surgeries. The open-source community, advanced in partnership of academic research and private concerns, offers solutions to a number of the issues faced, although no single software toolkit can be considered adequate to address all the issues foreseeable in the near future. Some such issues include integration of video sources, which occupy a central role in an increasing number of minimally invasive procedures, and the use of tools that leverage the parallel processing capabilities of modern GPUs.

The architectural design presented in this paper aims at creating a unique synergy among some of established open-source packages developed over the last decade. Using modern Computer Science principles, the current problems of intraoperative image-guided surgery can be addressed to yield the software architecture and the modular tools it hosts. This manuscript describes the architecture design principles and underlying components, followed by the modular functionalities that leverage them. The resulting functionality is demonstrated in three surgical scenarios (each in preclinical cadaver studies) with varying requirements and surgical tasks - specifically, skull-base, spine, and thoracic surgery.

2. DESIGN

The software architecture – referred to as TREK – is founded on two major complementary software packages: the cisst libraries for surgical navigation and the 3D Slicer application package for image visualization and analysis, each of which utilize other open-source software libraries such as VTK (Visualization Toolkit, KitWare, Clifton Park NY) and ITK (Insight Toolkit, Kitware) and leverage a broad base of expertise among the open-source community (Fig. 1).

![Figure 1](image-url) Illustration of the TREK software architecture for image-guided surgery system integration, showing cisst libraries are employed for navigation, device support, multitasking, and high-performance vector manipulation in combination with the 3D Slicer application for front-end visualization multi-modality image rendering and analysis.

The first major component incorporated within the architecture is the cisst package (Johns Hopkins University, Baltimore MD) – a collection of software libraries, specifically designed to ease the development of computer-assisted intervention systems. A number of surgical robotics systems ¹ employ cisst for core functionalities and real-time interaction with hardware devices ², with the developments described below adding to its functionality in the context of image-guided interventions. In the TREK architecture, cisst is utilized through use of its multi-tasking framework for simultaneous processing of external input, device support (most notably, surgical trackers), and vision library for grabbing and processing image/video data.
The second major package of TREK is the 3D Slicer application (Brigham and Women’s Hospital, Boston MA), an open-source software package for visualization and analysis of medical images. Slicer complements cisst by offering modules that can load medical images with editable planning information, a well-developed GUI toolkit, and the XML-based Medical Reality Modeling Language (MRML) which is used to create the virtual surgical scene.

These two software packages, their underlying components, and other external libraries and toolkits such as CUDA, SciPy, mlabwrap, SWIG, and OpenCV provide critical functionality and introduce a complex dependency challenge on the software side. The situation on the hardware side is also analogous, with each device using a different communication protocol and distinct accessibility through its custom application programming interface (API). In the TREK architecture, this problem was addressed using a high-level scripting language, and multi-level abstraction through object-oriented programming.

The use of scripting has been adopted by the academic community for some time, primarily with the popular use of the MATLAB computing language (Mathworks, Natick MA). The TREK architecture incorporates an alternative scripting platform – Python, an interpreted, general-purpose, high-level programming language – to bind together the underlying software packages while providing a high-level syntax to ease development. The dynamic nature of the language allows for loading of required components that enable more efficient use of limited computing resources, while the dynamic typing simplifies the task of abstraction.

One specific embodiment of this design is the trekVideo object. TREK pays special attention to video sources as many minimally invasive procedures employ them as the primary real-time feedback loop to the surgeon (Fig. 2) – the “eyes” of the minimally invasive surgeon. Specifically the use of video endoscopy in many minimally invasive procedures provide an access channel into the real-time imaging chain, enabling TREK to exploit the information content, and augment the video input with other sources of image or planning data. To this end, trekVideo facilitates simplified access to a number of video cameras via cisst, basic image processing, rectification, resizing operations through OpenCV, and visualization in the virtual scene using VTK and OpenGL graphics primitives. Camera extrinsics and the location of the image plane can be recorded in association with tracked tools – e.g., the tracked thoracoscope in Fig. 1. The resulting object is thus self-governed, continuously updating its image, pose and processed according to the defined filters, such as those responsible for resizing and correcting distortion. A number of these trekVideo objects can then be instantiated and are related to the VTK virtual scene through registration.

**Figure 2.** TREK setup showing example devices incorporated by TREK for preclinical studies in thoracic surgery. The C-arm arrangement in this case includes an infrared tracker (Polaris Vicra, NDI, Waterloo ON) aiding triplanar slice navigation. In addition, a tracked endoscope allows registration and fusion of the real-time video scene with CBCT and planning data.
3. RESULTS

The architecture and its impact upon surgical guidance capabilities can be appreciated first in terms of its underlying modules (Section 3.1), and second, in terms of a variety of task-specific embodiments for distinct surgical applications (Section 3.2), demonstrated below in relation to preclinical studies of CBCT-guided surgical performance.

3.1. TREK Modules for Surgical Guidance

i) 3D-3D Rigid and Deformable Image Registration. The preoperative CT can be registered to intraoperative CBCT using 3D-3D image-based deformable registration based on a fast implementation of a multi-scale variant of the Demons algorithm. The deformable registration also serves to bring preoperative surgical planning data to the updated context of the CBCT and, of course, any previous CBCT forward to the most up-to-date CBCT. Furthermore, any other preoperative image data (e.g., MR and PET) that has been registered to the preoperative CT may be registered to intraoperative CBCT via the same deformable transform. The method is implemented as a module in TREK, using ITK to carry out volumetric processing of the intensity values, followed by an iterative application of the core algorithm. The iterative intensity matching of CT-to-CBCT requires ~10-20 sec for registration of 256\(^3\) voxel CBCT volumes. The approach demonstrates significant improvement over rigid registration alone, with target registration error (TRE) ranging from 0.8–2.6 mm in initial cadaver studies. Modular implementation into TREK has been achieved using base algorithms and optimized volumetric processing offered through the ITK software library.

ii) Multi-Modality Tracking Configurations. Multi-modality trackers, including infrared (IR), electromagnetic (EM), and video-based trackers, can be initialized, configured, and controlled to track the patient and surgical tools. Novel arrangements are designed and tested using the central placement of the C-arm, as in the case of the “tracker-on-C” configuration shown in Fig. 1, where a video-based tracker is placed on the detector side of the C-arm, facing the isocenter and thereby providing both tracking as well as a useful video perspective on the surgical scene.

The trackers are rigidly registered to the scene using either a conventional fiducial-based rigid registration (Horn’s method) or an automatic method that leverages colocalization of fiducials identifiable by both the tracker and the C-arm images – for example, multi-modal markers that are visible both by trackers and in C-arm projections or CBCT (e.g., retroreflective spheres with a tungsten BB at the center). Online, real-time reporting of registration errors computed within the navigation system helps in guiding the operator in selecting the optimal combination of fiducials, with integrated model-based prediction of target registration errors for transparent visualization of the expected accuracy.

iii) Video-CBCT Registration / Augmentation. Image and tracker registration alone provides a conventional virtual representation of the surgical scene, where the updated information is presented through triplanar slice views or specific perspectives of a 3D view such that they do not occlude the target structures. Although such methods are adequate for some surgical scenarios, the increasing use of endoscopy/microscopy in minimally invasive procedures present additional, valuable display to the surgeon, which may augment or even supplant conventional (slice-based) guidance displays. The ability to register and overlay a wealth of multi-modality image and planning data directly within the endoscopic video scene presents an important solution to the problem of “too-much-data” in several surgical scenarios.

In TREK, this issue has been addressed through the creation of augmented reality tools that reference image and planning data directly to the video source. For example, a sinus endoscope of thoracoscope may be tracked via surgical trackers, and the video image is corrected for lens distortion through prior calibration. An additional step raises the possibility of higher levels of video-CBCT registration precision – specifically, an image-based fine-tuning of the registration that matches scale-invariant feature transforms in the video image to those in the surface rendering of bony structures in the CBCT, providing sub-mm overlay accuracy that surpasses that achieved using trackers alone.

iv) Real-Time DRR for 3D-2D Registration and Virtual Fluoroscopy. TREK enables digitally reconstructed radiograph (DRR) generation, through the use of the CUDA toolkit (NVIDIA, Santa Clara CA) contained in a module. GPU-accelerated forward projection methods (e.g., variations on the Siddon algorithm) have been implemented that are capable of creating virtual 2D fluoroscopic images from the latest acquired CBCT scan. The DRRs can either be used in an iterative fashion to register acquired fluoroscopic images to the 3D scene, or as a virtual fluoroscope that uses the latest scan to generate 2D images from oblique angles defined by a handheld tool. In TREK, the generated images of the latter
method are used as a video source handled by a trekVideo object, and the software uses a subset of the video augmentation module to achieve fluoroscopic images with superimposed planning data.

3.2. Translation to Specific Surgical Tasks and Applications

The surgical scenarios presented below emphasize the advantages of intraoperative imaging and, in particular, the streamlined workflow and visualization achieved by registration and fusion of image, planning, and video data via real-time video augmentation. The varying requirements of each surgical task call for a specific subset of the functionalities that exist in TREK, which can be achieved through specific configurations of its modular architecture.

i) Skull Base Surgery. Excision of tumors located at the skull base is often challenged by proximity to surrounding anatomy that commonly includes critical soft tissue structures, such as the carotid arteries and the optic nerves. Relying solely on conventional navigation methods (i.e. using preoperative image data alone) can be misleading and prone to geometric error. Soft-tissue anatomy is especially prone to change and deformation, even without surgeon interaction – for example, change in tumor morphology or sinus contents in the time interval between surgical planning and intervention. Even for transnasal skull base tumor surgery, deformation associated with herniation of the orbits (laminar papyracea) during intervention poses a significant limitation to conventional navigation. An intraoperative scan, followed by a deformable registration was shown to account for these changes and is readily integrated into the surgical workflow using TREK. However, even with an up-to-date image, registration with the tracking system may introduce errors. Considering the fiducial-based registration methods, the registration accuracy degrades as one moves away from the vicinity of the registration fiducials, e.g., as the surgeon progresses through the nasal cavity. The integrated video-registration is used to account for this discrepancy, giving an improvement in surgical precision beyond that achievable with tracking alone.

The reliance of video endoscopy in modern skull base surgery motivates the use of augmented reality techniques, thereby conveying a wealth of image and planning data in a naturally perceived, intuitive format and context of the video scene. Using VTK, the 3D rendered models of the surgical plan can be superimposed, their visibility adjusted, and a cut plane at the intersection of the endoscope image plane ensures that the bypassed structures are not displayed in the augmented video scene (Fig. 3).

ii) Spine Surgery. Procedures such as precise transpedicular intervention and vertebral reconstruction, where bone tissue visualization is of primary interest, are a further application for image-guided minimally invasive procedures. Rigid and/or deformable registration are applied in bringing preoperative CT (or MR registered to CT) to the up-to-date context of CBCT, with the latter playing an important role in capturing patient pose changes due to the articulated nature of the spine. In addition to the basic capabilities offered by TREK (as described above), the “Tracker-on-C” setup demonstrates a promising arrangement, allowing not only a tracking configuration with improved line-of-sight logistics but also a video view of the scene and tools that closely matches the surgeon’s perspective from above the table (Fig. 4). Recognizing fluoroscopy in combination with CBCT as the emerging modalities for high-precision spine surgery, the use of virtual fluoroscopy (DRRs) as mentioned above offers opportunities for streamlined workflow and reduced radiation dose to the patient and surgeon. TREK handles the
virtual fluoroscopy DRR as a video source, allowing the surgeon to quickly consider an approach from all angles and, equally importantly, providing augmentation of volumetric image and planning data directly within the virtual (or real) fluoroscopic scene naturally within the navigation architecture. Coupled with 3D/2D registration, virtual fluoroscopy can also be used to guide fast setup of the C-arm at tableside into an exact position that provides the desired field of view and collision avoidance, relieving the cumbersome process of trial-and-error (and spot radiographs from multiple perspectives) that is representative of conventional C-arm setup.

iii) Thoracic Surgery. For the task of guiding the surgeon to a sub-palpable nodule in a deflated lung, soft-tissue detail is essential. In this context, deformable registration is currently under development to account for the large deformation between the inflated and deflated lung. Thoracoscopic video-to-CBCT registration provides an intuitive visualization, as illustrated in Figs. 5 and 6. As shown in Fig. 6, the registration of preoperative imaging (CT), planning data (target and normal structure segmentations), intraoperative CBCT, and video (e.g., either the thoracoscope or Tracker-on-C), provides a useful view in assisting the placement of intercostal ports giving the best access to surgical targets. As shown in Fig. 5, fusion of multi-modality image and planning data within the thoracoscopic scene provides improved visualization of target and normal structures laying behind the visible surface anatomy provides a useful augmentation that could help to improve surgical targeting and avoidance of critical anatomy.

Figure 5. Augmented video scene from a thoracoscope, aiding in intercostal port placement. The cadaver is positioned in a lateral decubitus orientation for video-assisted trans-thoracic (VATS) access to the left lung. Planning data include skeletal anatomy, critical soft tissue, and potential target structures.

Figure 6. Screenshot of TREK showing runtime controls in the left panel, an augmented reality view of the thoracoscope image, with the tracked endoscope shown in slice views navigating with respect to an up-to-date CBCT scan.

4. CONCLUSIONS

The development of a new C-arm CBCT based surgical navigation architecture (TREK) has been reported, which addresses the challenges of system integration around intraoperative imaging. The integrated guidance system has the potential to overcome the inherent limitations of image guidance on preoperative images alone, such as changes in the anatomy due to tissue deformation and excision. Within the modular software architecture, TREK integrates the navigational tools that specifically emphasize the advantages of C-arm CBCT guidance, while allowing for task-specific embodiments in their application to various surgical procedures ranging from orthopaedic to head-and-neck and thoracic surgery. The architecture keeps customized solutions and code replication to a minimum by simultaneously leveraging...
the underlying cisst and 3D Slicer packages to accomplish tasks that are not originally supported by either, such as augmentation of video endoscopy with CBCT images and registered planning data.

The next stages of development will focus on workflow analysis and surgeon feedback, where both qualitative and quantitative methods will be used in assessing the added value and performance of these modular functionalities. A specific emphasis is placed on video-augmented guidance, as this approach is believed to offer the surgeon the most natural perspective and intuitive interface to the enormous mass of image and planning data—a key primary means of real-time visualization beyond that provided by conventional slice and 3D rendered views. Augmentation methods will be comparatively tested, including the use of alpha and color blending, superpositioning of 3D models versus projection to 2D, and a modular layered visualization that displays as much information content as desired, without occluding the endoscopic video stream.

ACKNOWLEDGEMENTS

This work was supported in part by the National Institutes of Health R01-CA-127444. The prototype C-arm was developed in academic-industry partnership with Siemens Healthcare (Erlangen, Germany). The authors extend sincere thanks to Dr. Peter Kazanzides and Mr. Anton Deguet (Department of Computer Science, Johns Hopkins University) for assistance with the cisst package. Preclinical studies were carried out in the Johns Hopkins Hospital Minimally Invasive Surgical Training Center (MISTC) with support from Dr. Michael Marohn and Ms. Sue Eller (Johns Hopkins Medical Institute).

REFERENCES