ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 37 (2007) S144–S151

Challenges in image-guided therapy system design

Simon DiMaio,[a,*] Tina Kapur,[a] Kevin Cleary,[b] Stephen Aylward,[c] Peter Kazanzides,[d] Kirby Vosburgh,[e] Randy Ellis,[f] James Duncan,[g] Keyvan Farahani,[h] Heinz Lemke,[i] Terry Peters,[j] William (Bill) Lorensen,[k] David Gobbi,[j] John Haller,[l] Laurence (Larry) Clarke,[h] Stephen Pizer,[m] Russell Taylor,[d] Robert Galloway Jr.,[n] Gabor Fichtinger,[d] Nobuhiko Hata,[a] Kimberly Lawson,[a] Clare Tempany,[a] Ron Kikinis,[a] and Ferenc Jolesz[a]

[a]*Brigham and Women's Hospital, 75 Francis St., Boston, MA 02115, USA*
[b]*Georgetown University, 2115 Wisconsin Avenue NW, Suite 603 Washington, DC 20057, USA*
[c]*Kitware Inc., 28 Corporate Drive, Clifton Park, NY 12065, USA*
[d]*Johns Hopkins University, B26 New Engineering Bldg, 3400 North Charles St, Baltimore, MD 21218, USA*
[e]*CIMIT, 65 Landsdowne Street, Suite 200, Cambridge, MA 02139, USA*
[f]*Queens University, Goodwin Hall, Kingston, Ontario, Canada K7L 3N6*
[g]*Yale University, 310 Cedar Street, BML 332 New Haven, CT 06511, USA*
[h]*National Cancer Institute, 6130 Executive Blvd. MSC 7412 Suite 6000 Bethesda, MD 20892, USA*
[i]*University of Southern California, Los Angeles, CA, USA*
[j]*Imaging Research Laboratories, Robarts Research Institute, London, Ontario, Canada N6A 5K8*
[k]*GE Global Research,1 Research Circle, Niskayuna, NY 12309, USA*
[l]*NIBIB, Democracy Plaza Two, 6707 Democracy Boulevard Bethesda, MD 20892, USA*
[m]*UNC, CB #3175, Sitterson Hall, Chapel Hill, NC 27599-3175, USA*
[n]*Vanderbilt University, Nashville, TN 37235, USA*

Received 1 February 2007; revised 13 April 2007; accepted 16 April 2007
Available online 24 April 2007

System development for image-guided therapy (IGT), or image-guided interventions (IGI), continues to be an area of active interest across academic and industry groups. This is an emerging field that is growing rapidly: major academic institutions and medical device manufacturers have produced IGT technologies that are in routine clinical use, dozens of high-impact publications are published in well regarded journals each year, and several small companies have successfully commercialized sophisticated IGT systems. In meetings between IGT investigators over the last two years, a consensus has emerged that several key areas must be addressed collaboratively by the community to reach the next level of impact and efficiency in IGT research and development to improve patient care. These meetings culminated in a two-day workshop that brought together several academic and industrial leaders in the field today. The goals of the workshop were to identify gaps in the engineering infrastructure available to IGT researchers, develop the role of research funding agencies and the recently established US-based National Center for Image Guided Therapy (NCIGT), and ultimately to facilitate the transfer of technology among research centers that are sponsored by the National Institutes of Health (NIH). Workshop discussions spanned many of the current challenges in the development and deployment of new IGT systems. Key challenges were identified in a number of areas, including: validation standards; workflows, use-cases, and application requirements; component reusability; and device interface standards. This report elaborates on these key points and proposes research challenges that are to be addressed by a joint effort between academic, industry, and NIH participants.

Introduction

The field of image-guided therapy (IGT) – sometimes also called image-guided intervention (IGI) or image-guided surgery (IGS) – has evolved from early stereotactic methods to modern multi-modal image-based navigation systems and has experienced many exciting advancements, particularly in the area of minimally invasive intervention. Much of the early innovation occurred within the field of neurosurgery, particularly for the treatment of brain tumors (Henderson and Bucholz, 1994; Bullitt et al., 2004). The nature and structure of the brain, and many of the tumors that invade it, create a frustrating compromise between tumor eradication and the sparing of functionally critical tissue (Claus et al., 2005). Modern image-guidance techniques improve the

* Corresponding author. Fax: +1 617 582 6033.
E-mail address: simond@bwh.harvard.edu (S. DiMaio).
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2007 Published by Elsevier Inc.
doi:10.1016/j.neuroimage.2007.04.026

visualization of pathologies with respect to adjacent tissue structures during tumor resection. They are used for precisely positioning and manipulating instruments and ablative devices. This integrated image-based approach has been adopted in many other clinical application areas and now involves advanced intra-operative imaging, image registration, image segmentation, visualization, navigation, and minimally invasive ablative therapies and robotics (for examples, see Shen et al., 2004; DiMaio et al., 2006, Peters, 2000).

The field of IGT system development has been advancing rapidly: major academic institutions and medical device manufacturers have produced IGT technologies that are in routine clinical use, dozens of high-impact publications are published in well regarded journals each year, and several small companies have successfully commercialized sophisticated IGT systems. In ad-hoc meetings held between several investigators in IGT over the last two years, a consensus emerged that to take the research and development effort in IGT systems to its next level of impact and efficiency a few key areas must be addressed collaboratively by the community. These meetings culminated in a two-day workshop that brought together several US-based and primarily NIH-funded academic leaders as well as industrial leaders in the field today, with discussions spanning many of the challenges currently faced in the development and deployment of new IGT systems. These challenges include identifying gaps in the engineering infrastructure available to IGT researchers, developing the role of research funding agencies and the recently established US-based National Center for Image Guided Therapy (NCIGT), and facilitating the transfer of technology among NIH-sponsored research centers. Four specific key challenges were identified in this meeting, namely: (1) How to increase the creation and exchange of reusable components; IGT systems are complex and not every group should have to construct a platform from the ground up. The tool development process needs to be made more efficient by leveraging and improving existing toolkits. (2) The need for performance standards for validation. We must have a common understanding of how to evaluate the performance of an IGT system and its components. A fundamental point that must be understood is that mission-critical software is evaluated not by its average performance but by its worst-case performance. (3) The need for increased awareness of the utility of use-cases and surgical/interventional workflows that is critical to building clinically acceptable IGT systems. (4) The need to motivate industrial partners to provide Application Program Interfaces (APIs) and research interfaces for their software/devices.

In the remainder of this report, we present a summary of the discussions that took place at the breakout sessions of the workshop on topics covering: Workflow, Validation, Tracking, and Robot Interfaces—identified by the authors as important areas for in-depth study of IGT system challenges (Technology focus areas), followed by a synthesis of the key research priorities that were identified in these discussions (Research priorities), and recommendations made by the participants for the role that the NIH (The role of the NIH in the development of IGT) and the NCIGT (The role of the NCIGT in the development of IGT) can play in the development of IGT systems in the future. While this paper focuses primarily on NIH-sponsored research in the United States, it is noted that there are many other funding agencies both domestically and internationally. We expect that many of the issues addressed in this report will be broadly relevant and applicable.

Technology focus areas

IGT workflow design

The science of workflow gained prominence in the 1970s as a tool to study the movement of documents in businesses. In a typical business setting, the goal of workflow analysis is to model document movement in such a way as to evaluate efficiency, quantify latency, and thereby, drive the allocation of resources. For example, in medical data management, the science of workflow is used to study the movement of patient records, procedure requests, insurance forms, and billings through hospitals.

More generally, the study of workflow is the analysis of task and resource scheduling: what tasks are needed to be performed, what resources are needed for each task, what orderings and synchronizations are needed between tasks, and how tasks are tracked. For image-guided therapies, workflow analysis has two primary applications. Workflow analysis can be applied to choreograph the movement of clinicians and technicians ("physician workflow") so as to reduce procedure time and patient risk (Paggetti et al., 2001). Workflow analysis can also be applied to study the movement of information and images within the computer that drives the image displays (data workflow) so as to speed processing and increase accuracy (Paggetti et al., 2001).

During workshop discussions, the concept of workflow was primarily focused on physician workflow. The rationale for this focus was that by understanding and quantifying physician workflow, developers will be better able to design and compare user interfaces and data workflows in IGT software (Dickhaus et al., 2004; Siddoway et al., 2007). For example, storyboarding – in this context – is the process of studying human–computer interactions by prototyping the user interface and its associated user interactions in a series of slides, such as in presentation software like PowerPoint. This is an outstanding means for expressing workflow and fostering communications between computer scientists, application developers, and clinicians.

This section describes highlights from our workshop discussions of the value of workflow, workflow analysis, and templates.

Workflow analysis and value

Workflow is an integral part of risk analysis and validation for IGT applications. Focusing on workflow aids the development of re-usable IGT libraries and applications and leads to the development of model-driven architectures. Therefore, our goal in software systems development is to create model-driven IGT libraries and applications that facilitate software review, test, reuse, and integration.

Methods for determining performance metrics, such as accuracy and time estimates during workflow simulation, as well as in the operating room, need to be developed. These methods will in turn need to be validated against measures acquired during phantom studies and actual procedures.

Workflow templates

The concept of a workflow template or model creates a framework in which applications can be developed or instantiated with specific algorithms that match the application's tasks. This modularity is inherent in the data workflow of one of the few research-grade open-source IGT software applications in use today, the Insight Toolkit (ITK), for example (ITK, 2007). Its utility for

IGT physician workflow for human–computer interactions was studied by Trevisan et al. (2003). He concluded that as few as four workflow templates are enough to model most image-guided surgery systems. From this it appears that Petri Net representations of workflow are frequently overly flexible and complex for most IGT applications and that the use of templates allows complexity to be appropriately managed.

The research challenge is to develop a theoretical and practical foundation for adapting workflow templates for a specific IGT application that is specialized to the clinical site, physician, and/or patient. This adaptation must ensure that options for problem solving and contingencies are not limited or overly constrained by the workflow template in the operating room during surgery.

Workflow execution models

Once workflow templates and adaptation mechanisms have been developed, it will become necessary to build a workflow execution model to translate workflow descriptions into functional data flows and user interfaces, as well as to enumerate and handle error conditions. The consensus among several developers of existing IGT toolkits and interfaces was that this execution model should be truly GUI and toolkit independent, cross-platform, and open-source, such that it can form a common basis for bridging existing IGT toolkits and application frameworks, including the major research-grade open-source IGT toolkits in use today; namely the 3D Slicer (3D Slicer, 2007), IGSTK (Gary et al., 2006; IGSTK, 2007), SIGN (SIGN, 2007), and a few others.

Validation of new IGT approaches

In general, system specifications are developed through a "requirements elicitation" process. However, clinical therapeutic tasks are complex and a new system design can typically only be characterized in limited ways. This has a significant impact on subsequent testing and validation, as system requirements and specifications serve as a natural baseline for evaluation. There is a tendency to equate greater precision with improved clinical outcomes, which is not always valid. Therefore, specifications may be too tight for a particular clinical need. In contrast, operator acceptance alone is too low a standard. After bench tests meet specification, new systems are typically evaluated in more realistic settings to determine:

- operating range,
- fault modes,
- tolerances, and
- peri-system compatibility.

The conundrum of specifications is that: prototypes and products are built to meet design goals, which are represented by specifications. In developing new techniques, there is an implicit assumption (which should be verified under use-testing, as described below) that meeting the specifications will create a tool or system that enables superior clinical results.

Here we explored two levels of system validation, namely *user evaluation* and *clinical outcome testing*.

Initial user evaluation

Comparative studies may be undertaken, successively, through retrospective analysis, simulators, phantoms, animal models, and human subjects. Present generations of simulators are insufficiently realistic to provide much assurance that a new device design is better than an old one for a complex task. Animal models provide much more realistic test conditions but suffer from the obvious differences in anatomy and physiology when serving as surrogates for humans; therefore, some level of human testing will be necessary.

Various groups are using techniques developed in other fields to characterize system performance. Several studies of simulators for laparoscopic surgery training have been conducted. More recently, tests have been made under actual OR conditions in animal or human models. For example, the Hager et al. at Johns Hopkins University have analyzed kinematic data recorded from the da Vinci system (Burschka et al., 2005), and the Vosburgh et al. at CIMIT/BWH have studied the performance kinematics and also the display utility in laparoscopic and endoscopic systems (Vosburgh et al., 2007).

At this level, various possible system error modes can be delineated and avoidance, mitigation, or response plans developed.

Clinical outcomes

The standard method for validating a new therapy is by evaluating its performance relative to standard practice. Almost always, a prospective clinical trial is necessary to validate a new approach. As examples of the level of effort that is traditionally required, consider the studies by Shapiro et al. (1989) for validating new methods for the treatment of hybrid astrocytoma. These took 5 years, and were well supported with a clinical infrastructure. In a Scottish study of 107 liver resections (Schindl et al., 2005), the fraction of liver tissue remaining after various procedures was measured. The study was helped by the fact that liver resections are very indicative of near-term outcomes.

In comparison to testing new surgical therapies, drug or vaccine trials have defined end points: markers or direct measurements such as tumor size. Controls may be easily implemented through placebos, which are much simpler than sham surgery. Drug trials are primarily interested in finding side effects; however, for surgical devices the standard has been lower. Surgical side effects (complications) are limited in number and are somewhat predictable.

Clinical outcomes are difficult to measure, and proper control groups are difficult to establish. It is often challenging to develop adequate patient numbers to give statistical power, particularly for identifying rare and unsafe conditions. Additionally, multi-site studies are needed for eventual FDA approval. This complexity may drive the adoption of a partitioned approach, in which anecdotal analysis is combined with statistically valid results on a lower dimensional space of factors. A model is then required to combine these dissimilar observations. Thus, as was stated: "one needs standard deviations but also the estimate of the number of dimensions." In addition, investigators will be well served to find creative ways to study multiple approaches simultaneously so that some level of serial analysis may be precluded.

Tracking and localization systems

In the context of image-guided intervention, the term "tracking" is a broad one that can include the act of localizing surgical instruments, therapy devices, patient anatomy, tissue targets, and even medical personnel as they move about the operating room. Workshop participants focused primarily on systems that track the

position and orientation of instruments and devices (Welch and Foxlin, 2002), for the purpose of establishing and maintaining a correspondence between medical images and the surgical field of view while navigating instruments during surgery. Our discussions highlighted challenges in two areas of interest, namely: (i) performance assessment and validation; and (ii) open systems and Application Programming Interfaces (APIs).

Assessment and validation

There are many ways to evaluate and report the performance of a tracking system, and testing methods are very much application-dependent (Nafis et al., 2006). Unfortunately, to date there has been no consensus on tracking requirements. Vendors report that they are reluctant to define requirements or standards, due to their exposure to liability, and the authors are not aware of any standards body that currently exists to govern performance specifications specifically for clinical tracking systems. As a result, it is difficult to compare systems based on their reported performance parameters. For example, typical performance metrics and measures include "average error" and "root mean squared error" with their associated standard deviation or confidence levels. These measures are of little use without knowledge of the testing procedures employed. For example, tracking accuracy will usually vary over the active workspace and depend upon the state of motion of the tracker. For electromagnetic trackers, one needs to further define the testing environment as magnetic distortions or electromagnetic interference can have significant impact on performance. Key technical performance criteria include: static accuracy, dynamic accuracy, static and dynamic precision, temporal resolution (i.e., update rate), spatio-temporal stability, latency, environmental sensitivity, interference between devices, and confidence reporting (the ability of the tracking system to "self-assess" and report the quality of its measurements).

Clearly, without standardization of testing methods, the combination of these criteria presents an intractable performance testing and specification problem. Testing methods for medical trackers should be based on clinical requirements and use cases since this is the context in which they will be operated. Unfortunately, clinical requirements are also difficult to determine as demands vary from medical procedure to procedure and from physician to physician.

Related to the problem of assessment and validation is the reporting of confidence measures by the tracker hardware during operation. In medical applications, it is important to have a continuous assessment of the quality of the measurement, with immediate notification of significant degradation. At present, some systems associate a confidence measure with tracked coordinates; however, these confidence measures are not consistent between vendors and are difficult to interpret quantitatively. Workshop participants felt that the availability of richer performance measures would be useful for developers. Industry participants indicated that, in many cases, such information is available within their systems, but can be extensive. Some dialogue between the scientific community, application developers, and device manufacturers is required to define the scope of this performance reporting, such that suitable data interfaces can be defined.

Open systems and APIs

Just as there is an absence of standards for assessing the performance of medical tracking systems, there are currently little or no software and hardware interface standards between vendors and devices. While each tracking system is different in its manner of operation, there is a need for a common API that can be used by software developers—this is particularly important in applications that integrate/fuse multiple tracking systems, and where some coordination or synchronization is required between systems (i.e., hybrid tracking).

The open-source model may be appropriate for helping to drive an "open interface standard" between devices, by giving vendors and developers a common software interface framework. There are a number of concerns with this model:

- interface requirements would need to be specified by determining a common set of functionality required by users and developers,
- regulatory approval and certification may be difficult to obtain; therefore, effective strategies for validating open software systems will be necessary,
- the deployment route through the open-source community is unclear, and
- the seat of responsibility/liability is unclear.

However, it should be noted that there is existing use of open-source software by vendors of medical devices (GEHealthcare-MicroCT, 2007; GEHealthcare-Specimen-MicroCT, 2007), and that this could serve as precedent. In such cases, open-source projects have been adopted and frozen for internal validation and deployment by vendors. An example of a promising open-source interface framework for tracking systems is the OpenTracker library (Reitmayr and Schmalstieg, 2001a,b; OpenTracker, 2007). Industry support for a common API will require some investment in time and resources. This means that vendors cannot be expected to support multiple APIs; therefore, it is necessary to build consensus between researchers and developers to support a single open-source interface, or at least a common specification of its requirements.

Interfaces to image-guided robots

Robots have assisted with surgery since the early 1990s, although currently their use is not as widespread as that of many other computer-assisted surgical technologies, such as the tracking and localization systems discussed above. However, it is clear that these technologies hold some important potential benefits for image-guided intervention, including:

- improved visualization and dexterity in areas that are difficult to reach, e.g., for minimally invasive surgery or for surgery inside CT/MR scanners,
- reduction of radiation exposure to surgeon, e.g., by removing the surgeon's hand from the fluoroscope field of view,
- provision of a "third hand", e.g., to hold cameras, retractors, etc.,
- increased accuracy in carrying out a surgical plan, e.g., the surgical equivalent of CAD/CAM; and the ability to work with smaller structures in microsurgical tasks, e.g., by motion scaling and/or tremor reduction, and
- improved safety via the use of virtual fixtures ("no fly" zones).

Workshop participants identified a number of key research, development, and deployment challenges in this area, namely: infrastructure for rapid prototyping, safety and validation, and control of commercial systems for research.

Infrastructure for rapid prototyping

The need for infrastructure support was raised by both industry and academia, though the specific needs are quite different. Manufacturers of surgical robots are interested in an infrastructure that would enable better technology transfer. This would include the ability to more rapidly integrate new technologies—such as those developed in academia—with their robots. Industry also expressed an interest in the software "best practices" that have evolved particularly in the open-source community (e.g., DART—the automated nightly testing framework initially developed for ITK) (DART, 2000).

Researchers expressed the need for an infrastructure to enable them to build robotic systems and applications to achieve their research goals. Significant hardware and software infrastructure is required to support research, particularly in IGT areas that involve medical imaging and navigation. Hardware support can include a number of different imaging systems (CT, MRI, X-ray, ultrasound, etc.) and several 3D tracking systems based on a variety of technologies (optical, electromagnetic, etc.). Software support includes standards such as DICOM, as well as open-source packages such as VTK, ITK, DCMTK, 3D Slicer, OpenTracker, and IGSTK. In contrast, there is no off-the-shelf robot system – with an open interface – that is suitable for medical use and no mature open-source packages for robot control.

Safety and validation

Several workshop participants raised issues about validation and regulatory approval, particularly in regard to the use of open-source software, such as how this software will be validated and who takes responsibility for maintenance. During the discussion, it was suggested that the best practice for medical device manufacturers wishing to use open-source software is to capture a "snapshot" of the software and validate their use of it as they would do for any third-party software. The manufacturer should apply its standard software change-control procedure and continue to use this version of software until it captures and validates a newer version.

This discussion also focused on the need for common phantom models that could be used to benchmark or validate systems being developed. This is a large effort due to the number of different target organs and surgical procedures that could be addressed by robotic systems. An ASTM working group (F04.05) is already developing a standard for measuring and reporting accuracy of computer-aided surgery systems; however, its initial focus is on the measurement accuracy of the underlying tracking technology (e.g., optical, electromagnetic, or mechanical system). Ultimately, we need phantom models that are more representative of clinical conditions since validation of clinical performance is paramount.

It was also noted that there is no standard for medical robot safety. This is a challenging area because safety requirements are very much application-dependent. In some applications, such as hip or knee replacement surgery, an occasional "glitch" of several millimeters may be tolerable, whereas in many other areas (e.g., brain surgery) this could be extremely hazardous.

Controlling commercial systems for research

Representatives from both US-based industry and academia agreed on the importance of bidirectional control of commercial systems for research purposes. This includes the need for integrating image feedback with robot systems. Therefore, it is not only important to have bidirectional control of commercial robots, but it is also important to have it for other devices such as intra-operative imaging systems.

The existence of external control functions requires careful validation, even if only intended for research purposes, because they must not compromise the performance of the device for its intended use. Clearly, there are safety and regulatory issues that must be resolved.

Results and discussion

From these technical focus areas, we have summarized a number of key research priorities for IGT systems development, as well as the role of funding agencies – such as the NIH – and the role of the NIH-funded National Center for Image Guided Therapy in catalyzing activity.

Research priorities

Requirements for IGT systems

Explicit performance requirements should be determined from the end users of these systems, i.e., the physicians and their medical personnel. Clinical needs may need to be interpreted by application developers to distill technical requirements; however, standards must come from the applications themselves. New methods are required for capturing and developing these requirements. In turn, common standards will help to drive – and make consistent – procedures for performance assessment and validation.

Hardware and software standards for IGT

Concerns raised by the FDA regarding the use of open-source software indicate that further discussions are necessary between industry, academia, and the FDA. Although some manufacturers have experience with open-source software, there is no "standard" procedure for incorporating this software. One possible outcome could be a FDA guidance document on the use of open-source software (as currently exists for the use of COTS software (FDA, 1999)). The dialogue should also include the topics of open architectures for, and bidirectional control of, medical devices.

Because devices such as tracking systems and interventional robots require so much specialized hardware, their use of open-source software may be more limited than in other fields, such as medical imaging. Nevertheless, even if a robot uses custom or proprietary software, the participants agreed that there is still great value in having open architectures and interface standards. This is also true for imaging devices, especially 2D and 3D ultrasound, which today have very limited research interfaces. This need for interfaces stems from the move toward more complex hybrid systems. In many cases, multiple standards do already exist; however, there is not enough agreement to facilitate and sustain collaborative development. There will always be competing standards; however, it is up to the marketplace which of these will prevail. Based on available precedents, it seems wise to allow "open-source" software technologies to be the driver of "open architecture" or "open innovation" trends in IGT, at least initially.

Some work is required to understand and develop the value proposition for industry to invest in opening interfaces and standards involving their devices. For example, therapy is a far smaller niche than diagnostic imaging today; therefore, it is not clear how to convince manufacturers/vendors of imaging systems to invest in new methods for which long-term pay-off is unclear,

particularly when significant engineering effort and cost are required to support standards. The current incentive to industry is that they can take advantage of resources and brain power that are being brought to bear by the research community; however, better matching between the research community and industry is required so that mutually beneficial progress is made. This requires a coordinated approach from the research community.

A further need within the research community is for greater compatibility between software toolkits for image guidance, with minimal duplication between toolkits, as far as it is possible. Following a number of established toolkits for visualization and image processing, such as VTK (VTK, 2007) and ITK (ITK, 2007), several efforts for building application software frameworks for IGT applications are already underway, including 3D Slicer (3D Slicer, 2007), IGSTK (IGSTK, 2007), and SIGN (SIGN, 2007). While it is unlikely that one single IGT toolkit will emerge for all applications, it would be helpful for us to align these efforts, to ensure optimal compatibility and interoperability.

Information and communications technology in IGT

Image-guided interventional systems typically consist of a number of components, devices, and software models that are connected through data and information interfaces of various forms. A number of these components have been developed in academic and industrial settings and in most cases exist as stand-alone systems with specific ad hoc proprietary or vendor interfaces. They can be considered as islands of IT engines and repositories with varying degrees of modularization and interconnection.

Information and Communications Technology (ICT) concepts have been studied for the purpose of mitigating the complexity of system integration across disparate interfaces. For example, the "Therapy Imaging and Model Management System" (TIMMS) is one attempt to deal with the information-intensive "Digital Operating Room" by complementing the image-centric world view of the classical PACS technology with an Information Technology model-centric view (Lemke and Vannier, 2006).

The collaborative development of highly modular systems will require that we develop ICT standards. An example of a relevant developing standard for medical imaging is that of DICOM, and the work of DICOM Working Group 24 (WG24). It is interesting to note that this is not only being driven by the traditional DICOM community, but that workgroups and committees now include surgeons, IGT engineers, etc. Therefore, the IGT community must be more active in this area to take advantage of the momentum that exists in the development of this standard. At the very least, if DICOM WG24 does not fulfill the basic requirements of IGT, then it would provide a good basis for initiating a workgroup for developing imaging and modeling standards for IGT, using ICT concepts and methodologies.

Consistent evaluation and validation methods

Between the cultural extremes of the bench engineer or scientist and the practicing clinician, we should build teams that can move us toward a unified philosophical approach and a mutually agreed representative paradigm for effective validation. This will not be static but rather will be improved over time.

An important first step will be to focus on developing effective methods to define requirement specifications for IGT systems as well as gold-standards. This will help to lead us to a consensus on how to evaluate and validate IGT systems from the low-level technical (e.g., tracking accuracy) to high-level clinical (e.g., survival/mortality rates). This clearly highlights the need for mechanisms to pool our results so that the community can converge on the most effective strategies and evaluation metrics. Note that the first objective of validation is to determine suitable measurements and metrics, while the second objective is to compare these metrics to the specified requirements. Without this context, the measurements are meaningless.

Knowledge databases and algorithm repositories

Workshop participants identified knowledge databases and shared repositories as means to address some of the difficulties in reaching consensus on IGT requirements, standards, and validation methods. Traditionally, academic journal publications have filled the need for sharing results and progress within research communities; however, the nature of IGT research and development means that we need to establish more extensive mechanisms for sharing and building upon progress in the field.

We need to develop algorithm repositories for open-source IGT software/hardware solutions, while leveraging existing toolkits to generate awareness of and access to existing algorithms. Open-source software is defined as being: (i) freely available to use, and (ii) distributed within a community of contributors. Therefore, the purpose of these repositories is to create awareness of existing technologies, as well as a forum for improvement and natural selection of superior approaches via the open-source mechanism. Similarly, repositories of IGT hardware design principles and knowledge repositories should also be considered.

The concept of "open data" is closely related to open-source, in the sense that it provides a context within which to compare and validate algorithms and methods. Image and data repositories have already been developed for medical image analysis (Holmes et al., 2005) any may serve as a template for the IGT community. For example, note the impact of the Fitzpatrick registration database (Fitzpatrick et al., 1998). To facilitate such data collection and dissemination, clinical researchers should be encouraged to design their IRB protocols to broaden the access to outcomes and data, so that the results can be used more widely and more effectively. Patient advocacy groups may also support mechanisms to make data and outcomes available for patients to share if they wish, just as it is possible for them to donate their blood and tissue.

Collection and dissemination of case studies and data can be time consuming and expensive, particularly if it is not part of standard clinical practice. Therefore, it may be helpful for the research community to develop uniform methods and tools for efficient data gathering with minimal overhead (Jalote-Parmar et al., 2006). As open-source systems become more widely adopted, such data gathering methods could be built in.

A library of reference workflows may also be of great benefit to the IGT community. For this, we will require standard tools and procedures for: (i) manually recording the actions of physicians and technicians in the OR (Fitzpatrick et al., 1998); (ii) for interviewing physicians to collect workflow descriptions for a variety of procedures; and (iii) for collecting data that are already automatically recorded by commercial IGT systems such as the StealthLink (Medtronic) and VectorVision (BrainLab).

The role of the NIH in the development of IGT

The NIH continues to play an important role in the support of research, development, and deployment of IGT technologies and systems in the United States. Based on the outline of current

research priorities presented above, the following activities were proposed as a means by which the NIH can further help to stimulate activity and collaboration within the community.

- Consider formation of a focused study section for IGT. To spark this action, it would be helpful to determine how many IGT-related proposals are currently entering the NIH review process. Such tracking is difficult due to the absence of codes for specific technologies as there are for diseases. Therefore, our priority is to define consistent keywords that can lead to codes for identifying IGT projects.
- Stipulate requirements for open-source software in NIH-issued RFAs and PAs to stimulate dissemination and sharing of methods and data. This will help to create an environment that is conducive to collaboration and consensus-building. Some care may need to be taken here to protect small business. Rather, emphasis should be placed on funding common open-source infrastructure that researchers and developers can leverage to add their own value and intellectual property. This will ensure that open-source infrastructure will be an enabler for technology start-ups, as well as the academic research community. What this community clearly lacks now is consensus and support for standards and infrastructure.
- Sponsored workshops should host open-source demonstrations and tutorials to create incentive and opportunity for small technology companies to become involved.
- NIH-issued program announcements and RFAs should be more explicit in their requirement for validation plans. In addition, it was felt that the NIH and the broader research and development community work together toward standard models for these plans and criteria for their evaluation.

The role of the NCIGT in the development of IGT

The National Center for Image Guided Therapy, sponsored by NCRR, NIBIB, and NCI at the NIH, is an important vehicle through which progress can be made in the research priorities listed above. Specific activities may include:

- The Center can be used to maintain and distribute consolidated knowledge databases of relevant open-source projects as well as case studies, benchmarks, performance metrics, and validation methods. To do this, the NCIGT can host workshops and symposia designed to stimulate discussion and convergence within the community. There is a possible role for NIST in establishing standards, as well as phantoms and procedures for validating IGT sub-systems. Therefore, such consolidated knowledge databases could be extremely helpful for building consensus on standards.
- Investigation and clarification of the scope of standards and regulatory bodies – such as the US Food and Drug Administration – in governing requirements for IGT systems and technologies.
- By fostering the development of software toolkits for image-guided systems (beyond VTK and ITK), the NCIGT should take the lead in developing functional specifications and by identifying the initial contributors to this effort, perhaps based on the knowledge repositories described above. Ultimately, these toolkits should define the standard interfaces between technologies.
- Finally, it is necessary for us to identify respected members of the IGT community who can champion and support methodologies and standards. These individuals should collectively represent the views of medicine, engineering, and research. The NCIGT is in a unique position to identify such champions, to bring them together in dialogue and to disseminate their views and recommendations through publication, reviews, and workshops.

Conclusion

The October 2006 workshop provided a forum for discussion between thought leaders in US-based academia, industry, and at the NIH on the state of the art in IGT systems engineering. Participants discussed current challenges in the development and deployment of new IGT systems, identified gaps in the engineering infrastructure available to IGT researchers, and provided recommendations to the research funding agencies at the NIH and the NCIGT. Four specific key challenges were identified in this meeting: (a) Increasing the creation and exchange of reusable components of IGT systems. (b) Developing a common understanding of performance standards and validation of IGT Systems and their components. (c) Increasing community awareness of the value of use-cases and surgical/interventional workflows for building IGT systems that are clinically acceptable. (d) Providing clear motivation to manufacturers to provide Application Program Interfaces (APIs) and research interfaces for their software/devices. Above all, participants strongly felt the importance of holding regular forums such as that reported here, to continue to refine the requirements for IGT as the technology develops, and to maintain an active dialogue between researchers and industry.

The challenges identified in this meeting span the breadth of IGT, stretching well beyond its early origins in neurosurgery, into many areas of contemporary therapy/surgery. Therefore, it is certain that research and development work aimed at addressing the research goals and challenges outlined in this paper, will have significant impact on the future of clinical practice.

Acknowledgments

This publication was made possible by Grants Number U41-RR019703 and U54-EB005149 from NIH. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

References

Bullitt, E., Jung, I., Muller, K., Gerig, G., Aylward, S., Joshi, S., Smith, K., Lin, W., Ewend, M., 2004. Determining malignancy of brain tumors by analysis of vessel shape. MICCAI Lect. Notes Comput. Sci. 3217, 645–653.

Burschka, D., Corso, J.J., Dewan, M., Lau, W.W., Li, M., Lin, H., Marayong, P., Ramey, N.A., Hager, G.D., Hoffman, B., Larkin, D., Hasser, C.J., 2005. Navigating inner space: 3-D assistance for minimally invasive surgery. Robot. and Autonomous Systems 52 (1), 5–26.

Claus, E.B., Horlacher, A., Hsu, L., Schwartz, R.B., Dello-Iaccono, D., Talos, I.F., Jolesz, F.A., Black, P.M., 2005. Survival rates in patients with low-grade glioma after intraoperative image guidance. Cancer 103 (6), 1227–1233.

DART, 2000. http://public.kitware.com/dashboard.php.

Dickhaus, CF., Burghart, C., Tempany, C., D'Amico, A., Haker, S., Kikinis, R., Woern, H., 2004. Workflow modeling and analysis of computer

guided prostate brachytherapy under MR imaging control. Stud. Health Technol. Inform. 98, 72–74.

DiMaio, S., Archip, P., Hata, N., Talos, I.F., Warfield, S.K., Majumdar, A., McDannold, N., Hynynen, K., Morrison, P.R., Wells, W. M., Kacher, D.F., Ellis, R.E., Golby, A.J., Black, P.M., Jolesz, F.A., Kikinis, R., 2006. Image-guided neurosurgery at Brigham and Women's Hospital. IEEE engineering in medicine and biology magazine: the quarterly magazine of the Engineering in Medicine & Biology Society 25 (5) 67–73.

FDA, Guidance for Off-the-Shelf Software Use in Medical Devices, US Food and Drug Administration, FOD #585, 9 September 1999.

Fitzpatrick, J., West, J., Maurer, J., 1998. Predicting error in rigid body. Point-Based Registration. IEEE Transactions on Medical Imaging 17, 694–702.

Gary, K., Ibanez, L., Aylward, S., Gobbi, D., Blake, M., Cleary, K., 2006. IGSTK: an open source software toolkit for image-guided surgery. IEEE Comput. Graph. Appl. 39 (4), 46–53.

GEHealthcare-MicroCT, 2007. http://www.gehealthcare.com/usen/fun_img/pcimaging/products/microctscanner.html.

GEHealthcare-Specimen-MicroCT, http://www.gehealthcare.com/usen/fun_img/pcimaging/products/specimicroct.html.

Henderson, J., Bucholz, R., 1994. Accurate and ergonomic method of registration for image-guided neurosurgery. Comput. Med. Imaging Graph. 18 (4), 273–277.

Holmes, D., Workman, E., Robb, R., 2005. Common access to uncommon data. Insight Journal. Dspace handle: http://hdl.handle.net/1926/22.

IGSTK, 2007. http://www.igstk.org/.

ITK, 2007. http://www.itk.org/.

Jalote-Parmar, A., Pattynama, P.M.T., Goossens, R.H.M., Freudenthal, A., Samset, E., De Ridder, H., 2006. Exploring a user centric methodology to investigate and integrate information gathered during medical intervention. In Proceedings of the 16th World Congress on Ergonomics, Maastricht, The Netherlands: International Ergonomics Association.

Lemke, H.U., Vannier, M.W., 2006. The operating room and the need for an IT infrastructure and standards. Int. J. CARS 1 (3), 117–121.

Nafis, C., Jensen, J., Beauregard, L., Anderson, P., 2006. Method for estimating dynamic EM tracking accuracy of surgical navigation tools. Medical imaging: visualization, image-guided procedures, and display. In: Cleary, K.R., Galloway Jr., R.L. (Eds.), In Proceedings of the SPIE 6141, pp. 152–167.

OpenTracker, 2007. http://studierstube.icg.tu-graz.ac.at/opentracker/.

Paggetti, C., Martelli, S., Nofrini, L., Vendruscolo, P., 2001. Interface design and evaluation for CAS systems. Proceedings of the 4th international Conference on Medical Image Computing and Computer-Assisted Intervention.

Peters, T., 2000. Image-guided surgery: from X-rays to virtual reality. Comput. Methods Biomech. Biomed. 25–27.

Reitmayr, G., Schmalstieg, D., 2001a. An open software architecture for virtual reality interaction. Proceedings of the ACM Symposium on Virtual Reality Software and Technology, pp. 47–54.

Reitmayr, G., Schmalstieg, D., 2001b. OpenTracker—An open software architecture for reconfigurable tracking based on XML. IEEE Virtual Real. 285.

Schindl, M., Redhead, D., Fearon, K., Garden, O., Wigmore, S., 2005. The value of residual liver volume as a predictor of hepatic dysfunction and infection after major liver resection. Gut. Edinburgh Liver Surgery and Transplantation Experimental Research Group (eLISTER) 54 (2) 289–296.

Shapiro, W., Green, S., Burger, P., Mahaley, M., Selker, R., VanGilder, J., Robertson, J., Ransohoff, J., Mealey Jr., J., Strike, T.A., et al., 1989. Randomized trial of three chemotherapy regimens and two radiotherapy regimens and two radiotherapy regimens in postoperative treatment of malignant glioma. Brain Tumor Cooperative Group Trial 8001. J. Neurosurg. 71 (1), 1–9.

Shen, D., Lao, Z., Zeng, J., Zhang, W.I.A., Sesterhenn, I.A., Sun, L., Moul, J.W., Herskovits, E.H., Fichtinger, G., Davatzikos, C., 2004. Optimization of biopsy strategy by a statistical atlas of prostate cancer distribution. Med. Image Anal. 8 (2), 139–150.

Siddoway, D., Ingeholm, M.L., Burgert, O., Neumuth, T., Watsona, V., Cleary, K., 2007. Workflow in interventional radiology: nerve blocks and facet blocks. In Proceedings of SPIE Medical Imaging 2006: PACS and Imaging Informatics 6145, 61450B.

SIGN, 2007. http://www.ncigt.org/sign/index.html. Trevisan DG., Raftopoulos, C., Macq, B., Vanderdonckt, J., 2002. Modeling augmented reality system, image guided surgery case study. Proceedings of the 1st Glasgow Context Group Colloquium GCG.

Trevisan, D., Vanderdonckt, J., Macq, B., Raftopoulos, C., 2003. Modeling interaction for image-guided procedures. Proceedings of SPIE Medical Imaging, vol. 5029, pp. 108–118.

Vosburgh, K.G., Stylopoulos, N., San Jose Estepar, R., Ellis, R.E., Samset, E., Thompson, C.C., 2007. EUS with CT improves efficiency and structure identification over conventional EUS. Gastrointest. Endosc. 65 (6), 866–870.

VTK, 2007. http://www.vtk.org/.

Welch, G., Foxlin, E., 2002. Motion tracking: no silver bullet, but a respectable arsenal. IEEE Comput. Graph. Appl. 22 (6), 24–38.

3D Slicer, 2007.http://www.slicer.org/.