Image Guided Surgical Interventions

Surgeons have traditionally performed procedures to treat diseases by gaining direct access to the internal structures involved, and using direct visual inspection to diagnose and treat the defects. Much effort has gone into identifying the most appropriate incisions and approaches to enable full access inside body cavities, specific organs, or musculoskeletal structures. Imaging has traditionally been used primarily for preoperative diagnosis and at times for surgical planning. Intraoperative imaging, when used, was meant to provide further diagnostic information or to assess adequacy of repair. In most cases radiograph static images or fluoroscopy have been used in the operating room. As the application of less invasive procedures has progressed, other imaging technology has been applied in an effort to address the limitations of simple radiographs or fluoroscopy. Computed tomography (CT), magnetic resonance imaging (MRI), ultrasound, nuclear radiographic imaging, and modified optical imaging have been introduced to provide the information required to plan and perform complex interventions inside the body without the need for direct open visual inspection.

In parallel with the developments in imaging modalities, endoscopic surgery has advanced with the introduction of rigid and flexible scopes equipped with video cameras to magnify and display the image obtained. Advances in optics and digital electronics have combined to provide unparalleled image quality even with small diameter scopes, resulting in an explosion of endoscopic procedures involving virtually every structure in the body. The only real limitation to imaging has been the inability to see or "image" through opaque structures, since the irradiating or "illuminating" energy provided through the scope has been almost exclusively visible light. This limitation has confined endoscopic surgery to areas where a natural body cavity or physical space can be accessed with a scope and instruments, and filled with nonopaque medium such as a gas or clear fluid. Despite these limitations, optical endoscopy has revolutionized the way many surgical procedures are performed, and has spawned a whole industry of instrument manufacturers that, in conjunc-

tion with surgeons, design task-specific tools capable of entering the desired body cavity alongside the scope. It has also forced surgeons to develop new techniques utilizing these specialized instruments for carrying out a wide variety of operative procedures.

For most endoscopic procedures, optical imaging provides sufficient information to guide the various steps of the surgical procedure including: identification of target lesion, instrument navigation from entry point into the body to the target, identification of important anatomic structures, view of tissue-instrument interaction in real-time, and real-time assessment of tissue reconstruction/manipulation.

Along with the focus on imaging and instrument development, there is widespread recognition that there are major components common to all image-guided procedures that have had to develop in parallel. These include: imaging modalities that provide the required information for successful achievement of the procedure; image displays with optimal resolution and ergonomics to facilitate workflow; aids to instrument navigation, particularly if the instrument, tissue target, or instrument-target interaction are not directly visible; and specialized surgical instruments and devices optimized for the task and modified for the type of imaging modality.

This monograph covers each of these 4 topics (imaging, displays, navigation, and surgical instruments) as they apply to the various imaging modalities. Our goal is to provide an overview of current technology available or being developed with examples of clinical applications and areas of research and technology development. The importance of this field to practicing surgeons cannot be overstated, since technological advances will likely enable performance of ever more complex procedures inside the body using minimal access techniques to limit collateral tissue injury and improve overall surgical results. Thus, for practicing surgeons, gaining an understanding of the basic concepts of this technology will become imperative to keep pace with innovations in minimally invasive surgery.

Imaging Technology

To obtain critical information regarding body structures for performing surgical procedures, beyond what is visible externally or by direct internal inspection, various forms of imaging, including optical imaging, are often required. The advantages of visible light endoscopy are high spatial resolution (limited only by the optics and digital camera resolution, on the order a few hundred microns), real-time imaging (ie, video rate, or 30 to 70 frames per second or more), and the ability to see not only the inner surface of cavities, but also structures such as instruments within the lumen of the

TABLE 1. Common imaging modalities

Modality	Spatial resolution	Tissue penetration	Differentiate tissue boundaries
X-ray fluoroscopy	++	++++ Complete	++
CT	+++	++++ Complete	+++
MRI	+++	++++ Complete	++++
Ultrasound	+	+++ No bone; 2 to 20 cm	+++
Optical	++++	++ 5 mm or less	++++

CT, computed tomography; MRI, magnetic resonance imaging.

cavity. The main limitation of optical endoscopy, however, is the lack of penetration of the illuminating energy (visible light) into solid or opaque tissue structures.

Traditionally radiograph imaging has been used as an adjunct or instead of optical imaging inside the body due to its ready availability, and several advantages over other imaging modalities including high spatial resolution, ability to penetrate all tissues, low cost, and its ability to guide interventions in real-time, with and without the aid of contrast material. The disadvantages of radiograph imaging, however, include poor differentiation between soft tissue structures, ionizing irradiation, and 2-dimensional (2D) projection imaging, the latter being a significant limitation to navigating instruments within body cavities. The various other imaging technology options available to surgeons can be categorized on the basis of specific factors that may be of advantage for a particular procedure such as: spatial resolution, defined as the ability to detect point targets as distinct (depends on the imaging path to the target); degree of tissue penetration; ability to differentiate tissue boundaries (image contrast); ability to detect different tissue characteristics (density, water content, mechanical properties, deformability); location of imaging energy source (tool tip vs. external); availability of real-time imaging; and use of ionizing versus nonionizing radiation to obtain the images.

Table 1 summarizes the common imaging modalities used in image guided procedures and categorizes them according to the factors listed above.

For real-time guidance of surgical procedures, imaging must provide information regarding the location of the target lesion, associated and nearby structures, the relationship of surgical instruments to the target, and do this in "real-time" or at least 25 frames per second. The capability of an imaging modality to provide this information depends ultimately on the physics involved in generating the image. First, tissue must be interrogated by some form of energy; the resultant signal must be minimized of noise and converted to an image. To appreciate how each imaging technique can be used in surgery, it is important to understand

Soft tissue characteristics	Energy source	Real-time	Ionizing radiation
++ Semi-transparent	External	Yes	Yes
++	External	±	Yes
+++	External	±	No
+++	External/tool tip	Yes	No
++	Tool tip	Yes	No

how the image is generated and what the advantages and limitations are for each technique.

X-ray Fluoroscopy

X-ray electromagnetic radiation is of higher energy than visible light and, as such, it penetrates bone and soft tissue. X-rays have wavelengths in the range of 10 to 0.01 nm. A photon originating outside the body may pass straight through the body or it may be absorbed or scattered by tissue. Variations in absorption by tissues of differing material densities are what is detected and presented in the resulting 2D image. Scattering is an unwanted effect whereby the x-ray is deflected from its original path. Scattered photons travel in many directions, but some of these ultimately reach the film or detector and blur the image. Real-time updating of generated images from flat-panel semiconductor detectors (fluoroscopy) is straightforward since X-rays travel at the speed of light.

This oldest of imaging modalities is still the one in most common use today in the interventional laboratory. It has been used to identify bony and sometimes soft tissue structures or, with the use of radio-opaque contrast material, to delineate cavities or blood vessel lumen. In the past, when film was used for image display, it required time for film development and analysis, often increasing substantially the length of the procedure. Now, with detectors capable of working at fast frame rates (>30 Hz), x-ray fluoroscopy is used to provide real-time imaging when images of rapidly moving structures are required.

The main drawback of x-ray, however, is the absence of 3-dimensional (3D) information, since it is fundamentally projection imaging, that is, a point source of energy that penetrates all structures to varying degrees and displays this information on a 2D image or film plate. Indirect 3D information can be obtained by generating orthogonal views such as with a PA and lateral chest radiograph; however, the surgeon is required to integrate the 2 projections in his mind to get a sense of the 3D spatial

localization. As the structures that are being imaged become more complex in geometry, this approach becomes more cumbersome.

Although fluoroscopy has the advantage of adding the time dimension required for some diagnostic and interventional procedures, it still does not provide information about different layers in soft tissue. Despite this limitation, for many open procedures, fluoroscopy is able to provide sufficient complementary information required by the surgeon. However, as interventions to treat defects have adopted less invasive or more remote access, such as balloon angioplasty of a vessel obstruction, additional information regarding soft tissue layers, physical relationships between key structures, and more precise 3D localization is required to carry out the procedures.

Since photons are projected through 3D tissue onto a 2D detector, a point in the 2D image represents the absorption along the entire 3D path length. As a result, image contrast is limited and details of fine scale tissue structure are lost. For the same reason, tissue boundaries, such as that of a blood-filled cavity, can also be indistinct. To improve boundary visibility, injectable intravenous contrast agents can be administered and mixed with blood to attenuate x-ray radiation over blood alone. The difference between blood concentrated with agent and muscle enhances image contrast and boundary detection—if only in 2D.

Computed Tomography

Since CT is generated by x-ray, principles regarding energy sources, ionization, and tissue penetration are the same as for x-ray fluoroscopy. What is different is the handling of the beam and mathematical reconstruction to the point of localizing a target in 3D space. Tomography refers to imaging by sections ($\tau o \mu o \sigma$ is Greek for "section") and so results in sharper images by focusing the x-ray beam at a level corresponding to a depth of interest. Rather than allowing beams to pass through without regard for depth or direction of travel, tomograms exploit changing the direction of travel to provide additional information about specific regions of interest. Older moving tomograms, where the x-ray source and detector moved rapidly in a circular motion focusing on 1 point in space, are historical examples of this technique, but modern gantries that rotate a beam are the physical incarnation of modern tomographic imaging. The second innovation enabling higher resolution tomography is mathematical reconstruction such as filtered back projection or iterative reconstruction. Although x-ray is the most common application of tomographic reconstruction, these computer techniques are also used in MRI, nuclear medicine, and ultrasound.

Tomographic images can be stored in slices as 2D pixels, but since the location of each slice is known, these data can be converted to volume elements (3D data) or voxels. Resolution, boundary detection, and image contrast, especially with use of intravenous agents, can be quite high, particularly between tissues of differing densities such as bone versus muscle. Modern computing allows sophisticated processing and reconstruction. Newer parallel "multislice" CT employs imaging of multiple slices simultaneously. This does not necessarily reduce radiation exposure, but it quickens acquisition time.

Although real-time 3D CT imaging is limited by the need to reconstruct the image from a large volume of data, advances in computing speed have nearly eliminated this as a technological constraint. Current CT systems, however, were not designed with the purpose of real-time imaging in mind. Most importantly, radiation dosing over the course of a surgical procedure can be prohibitive and appreciation of techniques that employ nonionizing radiation becomes significant.

Magnetic Resonance Imaging

MRI comes from the physical principle of nuclear magnetic resonance. The term "nuclear" is not synonymous with ionizing radiation; rather it refers to the resonant behavior of atomic nuclei when excited by radiofrequency energy in a very high magnetic field. Materials of this type, such as hydrogen, are referred to as paramagnetic. Their nuclei normally have arbitrary orientations with regard to their quantum "spin." These nuclei line up or orient uniformly in high magnetic fields. When excited by pulsed radio waves, some of these atoms temporarily move off alignment into a higher energy state. As they realign, they emit radio waves that can be detected. This emission can be localized by using a magnetic field with a known gradient in space. Ultimately, a 3D data set comprised of voxels can be synthesized. The electromagnetic radiation in MRI is not ionizing.

One of the fortunate occurrences in nature is that one of the most abundant atoms in soft tissue is the paramagnetic atom hydrogen. Since hydrogen is so abundant, especially in water, it provides an ideal substrate for this modality. Resolution can be improved using higher strength fields with higher gradients. As such, MRI can provide images of high soft tissue quality or contrast. As with x-ray, the use of intravenous substances in blood that modulate signal intensity can enhance resolution and boundary detection even further. Imaging of moving structures is possible by MRI but has been traditionally done, for the heart, for example, by recombining multiple beats "after-the-fact" by a technique known as

gating. *Gating* refers to controlling the signal acquisition using a detectable triggering event. In the heart the electrocardiogram is used as the triggering event since it is relatively regular and easily detectable. The ability to trigger acquisition at a particular time in the cardiac cycle enables summation of several acquisitions, which significantly improves the signal-to-noise ratio, since noise is random whereas the signal from the tissue is not. Since the radiofrequency signal obtained from tissue is weak, summation of multiple signal acquisitions enhances the signal and therefore the image resolution. For moving structures such as the heart, gating allows summation to occur from tissue at the same location. For real-time imaging, however, the number of cardiac cycles that can be summed is limited, since images must be refreshed at a sufficiently fast frame rate to allow the surgeon to navigate instruments inside the body in a reasonable time frame. Furthermore, instrument interaction with tissue causes tissue displacement producing cycle-to-cycle variation that may not be detected at slow frame rates. Live imaging with MRI has become possible but can have serious adverse effects on image quality when compared with gated techniques.

The most important limitations of MRI for procedural imaging, however, are the presence of intense magnetic fields and the presence of a large instrument usually surrounding the patient. Special imaging suites have been created as well as tools to work safely within this environment, but the image resolution and frame rate are still limited.

Ultrasound

Ultrasound is nonionizing and is unique among other imaging modalities in that it depends, not on the various forms of electromagnetic radiation, but on the vibration of physical media. Ultrasound is defined as acoustic waves with a pitch or frequency above the limit of human hearing; this is classically defined as 20,000 cycles per second or Hertz. Other animals such as dogs and dolphins can hear 40,000 Hz and higher. Imaging ultrasound employs sound waves that are inaudible by any animal—at least 2 million Hz or MHz. It is a fundamental property that waves traveling through a media are partially reflected when they encounter a change in the media's resistance to their travel (impedance)—as at the boundary between tissue structures. Thus, ultrasound waves typically are generated and received at the same location since they detect changes in reflection rather than absorption. Ultrasound energy of lower frequencies has better penetration and waves of higher frequencies (and shorter wavelengths) have better resolution. Thus, 2- to 4-MHz transducers are used to image the heart through the chest wall at 12 cm

depth and can only resolve points 1 to 2 mm apart, whereas small catheter-tip-mounted 30-MHz transducers are used to image inside a coronary artery, over a length scale of 5 to 6 mm, and with a resolution of less than 1 mm. At 30 MHz, the histopathology of an atheromatous plaque can be seen.

Two-dimensional ultrasound images are composed by stepping a focused beam (a scan line) across a plane and assembling each scan line's reflection to form the image. Three-dimensional images are composed by scanning adjacent planes. Imaging resolution depends on the spacing between scan lines and between scan planes as well as the echo quality of each scan line. Factors that reduce echo quality and lead to imaging artifacts include inhomogeneities in tissue (eg, chest wall fascia), multipath reflections (eg, at tissue-gas interfaces and in metal instruments), and off-axis reflections from highly reflective surfaces (eg, the diaphragm and urinary bladder). Newer transducers that have broader frequency spectra (for more penetration and better resolution), better beam formers with newer algorithms (using more information from each scan line), and harmonic imaging (detecting echoes at multiples of the input frequency) have substantially improved image quality over the past decade.

Ultimately, however, resolution depends on scan line spacing. In comparison with electromagnetic radiation, sound travels slowly, only 1540 meters/sec in water, and as more scan lines are used to generate an image, more time must be allocated to wait for returning pulses. This reduces the frame rate. Thus, there is a tradeoff imposed by the physics. A given frame rate fixes the maximum number of scan lines. One can increase scanning area by spreading lines farther apart, but this degrades resolution. Conversely, resolution can be enhanced by spacing lines more closely together; however, this reduces the extent of the scanning area.

Despite this physical limitation, modern clinical ultrasound systems are the only modality that provides 3D real-time images from a small portable system using nonionizing energy while not requiring special environmental considerations.

Multimodal Imaging

Multimodal imaging is a method of augmenting image guidance by combining different modalities into a single image. The goal is to provide the surgeon with complementary information about the anatomy that will serve to improve navigation and/or target selection. Examples of this type of image guidance include CT or MRI image overlays on the operative field for neurosurgery, and combined radiograph fluoroscopy and CT images of the left atrium and pulmonary veins for arrhythmia ablation. To

generate such images, the precise location of specific anatomic structures must be determined and aligned in all the images to combine them accurately. This requires creation of a coordinate system or frame for each imaging technique and then aligning the coordinate frames of all the image sets to produce the multimodal image. Consideration must also be made of the optimal way to display the data from each imaging technique so that the benefits of each modality are realized and it is also easily understandable by the surgeon.

The process of aligning the coordinate frames and anatomical structures of each modality is termed image registration. A sizable literature exists on this topic.[1] Alignment methods can be separated into 2 groups: fiducial-based and anatomy-based. Fiducial-based image registration uses passive manmade objects, termed fiducials, that are positioned on the body and easily recognized by the imaging techniques for alignment. Several of these fiducials are used to ensure proper image alignment. The fiducials must be rigidly fixed to the patient, which can be quite cumbersome. One method, used for neurosurgical imaging, involves use of stereotactic frames.[2,3] A frame is fixed to a patient's skull before imaging and appears in all nonultrasound images. Another method is to use small point fiducial markers fixed to the skin or the skull. These are easier to apply than a stereotactic frame and are suitable for noncranial anatomy.[4,5]

To enable real-time image registration for surgical guidance, fiducial registration can be augmented by use of a tracking system. Tracking systems (eg, electromagnetic, optical, mechanical) enable the calculation of the precise physical location in space of the images' fiducials in the tracking system's coordinate frame. Each modality's image is then aligned to the tracking system's coordinate frame by corresponding the image coordinates of fiducials with their physical coordinates, as measured by the tracking system. In addition to providing real-time image registration, the physical location of surgical tools can also be tracked with this approach and their location can be overlaid on the image, greatly aiding instrument navigation inside the body.

Anatomy-based image registration is an active area of research.[1] In this case, registration is also achieved by corresponding points between physical space and image coordinates, but anatomical features are used rather than artificial ones. Features must be segmented from the various imaging techniques to be used for the image alignment. Typically, anatomic structures that are easily imaged by all the imaging forms must be used. This technique can also be used for 3D reconstruction from multiple sequential 2D planes. In the case of 2D ultrasound, multiple
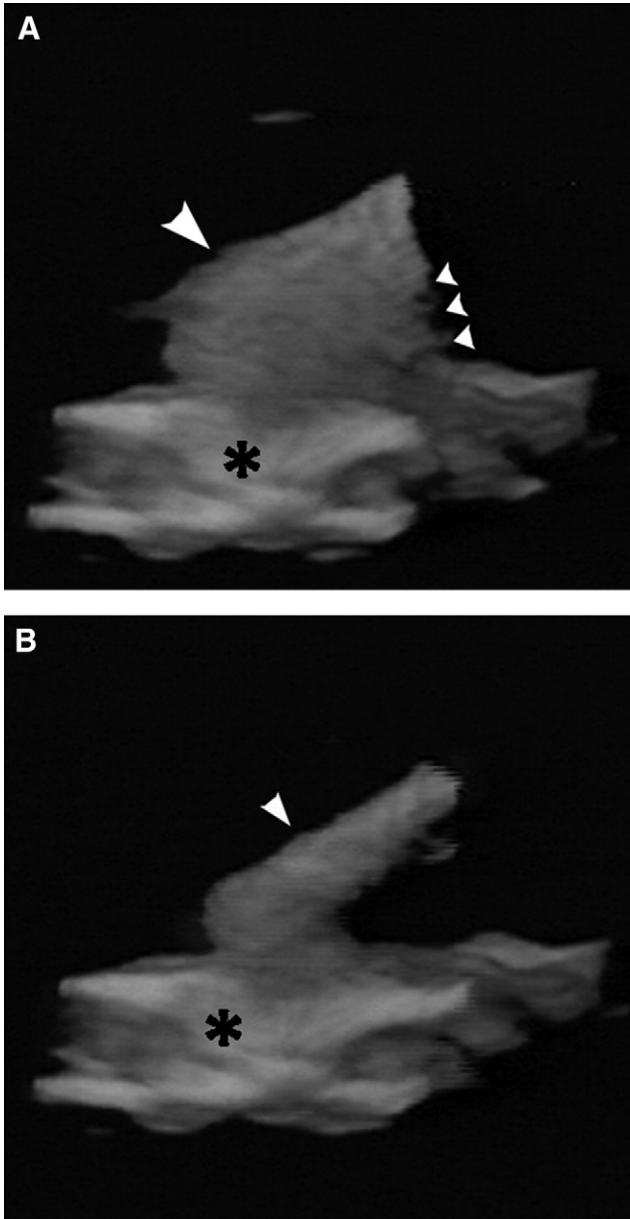
FIG 1. Side lobe artifact within real-time three-dimensional ultrasound images of a surgical tool (large arrowhead) approaching heart tissue (∗). These images show the wide variation in severity of side lobe clutter (small arrowheads) depending on material properties of the tool. A metal instrument (A) has a much greater impedance mismatch with the soft tissue than a graphite composite instrument (B).

images can be formed into a 3D volume for more accurate registration.[6-8] Suitable anatomy for image registration includes the surface of the spine[9,10] and hepatic vascular structures,[11] among others.

Navigation of Surgical Instruments

Instrument navigation refers to the ability to introduce or advance a surgical tool, catheter, or needle to a desired target within the body, usually with the surgeon choosing the trajectory. Often the trajectory is dictated by the location of the target region with respect to other critical anatomic structures and whether there is a readily accessible route to that structure such as a blood vessel, bowel lumen, or body cavity. The location of the instrument or tool with respect to anatomic structures is most commonly obtained by the imaging modality used to guide the intervention, but other methods of localization can also be used such as electromagnetic or infrared tracking of the tool tip. To date, a great deal of work has been published concerning image-based needle or instrument navigation using fluoroscopy, CT, and MRI, primarily for interventional radiologic procedures.[12-14] For surgical procedures, due to the need for real-time imaging to guide instrument navigation, radiograph fluoroscopy and ultrasound are two frequently used imaging modalities. With its ease of use and nonionizing energy for imaging, diagnostic and therapeutic navigation with ultrasound is a standard part of modern medicine.[15]

Ultrasound-based navigation poses two challenges, however. The first is ultrasound's limited field of view. When anatomical details within an image are unfamiliar or unrecognizable the clinician can lose orientation and is unable to mentally align the current image within the anatomy. In the case of 2D imaging, images are often taken parallel to image planes commonly displayed in reference texts (ie, axial, sagittal, and coronal). In many ultrasound applications, however, either acoustic windows do not allow acquisition in these planes or probe motion is constrained such that the planes cannot be achieved. In these cases, it is difficult to avoid disorientation. This is particularly true with 2D images and when the ultrasound transducer is not directly visible, as is the case with laparoscopic ultrasound.[16] The second challenge with ultrasound-based navigation is due to the distortions of the image by metal surgical instruments. These make it difficult for the clinician to know the precise relative position and orientation of instruments with respect to the tissue, which leads to uncertainty in the required instrument motions to achieve a task. Imaging artifacts arising from reverberation and side lobe reflections[17] can obscure tissue near the instrument and make it nearly impossible to identify the instrument tip in the image (Fig 1). Also, mirror-like
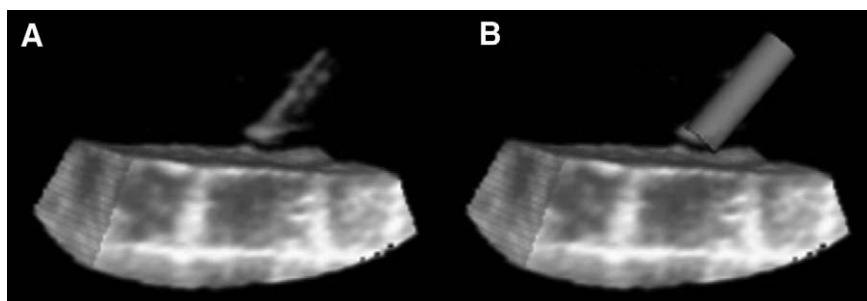
FIG 2. Ultrasound image of acetyl rod instrument. (A) Ultrasound image with side lobe artifact on the tip of the rod. (B) Image produced by image based tracking of the rod and creating a cartoon overlay of the actual instrument in the field of view. (Reprinted with kind permission of Springer Science + Business Media from Novotny et al.[24])

reflections from the instrument's surface can cause it to appear and disappear based on the angle between the incoming sound and the instrument's surface. In combination with boundary smoothing, these effects can make fine navigation difficult because instrument contact with tissue is unclear, potentially leading to tissue damage. Instrument tracking is one technique that can be used to alleviate this issue. Other techniques such as instrument surface modification are described below.

Instrument Tracking

Instrument tracking is useful to the surgeon when the surgical instrument is outside the field of view, when it is difficult to identify in the image due to the artifact that it produces, or when the instrument cannot be readily detected by the imaging modality. With optical or radiograph-based imaging (including CT), the latter is rarely the case since most surgical instruments are or can easily be made opaque to the imaging energy, light, or x-ray. However, for ultrasound or MRI, instruments are often difficult to identify in the operative field of view and tracking instrument position to enable display of the surgical tool in the operative field is a key component of instrument navigation. Thus, instrument tracking consists of determining the relationship between the instrument, the anatomy, and the image coordinate frame. Applications of instrument tracking include automatic instrument control and improved image displays. Two methods exist for automatically determining this relationship: external and image-based tracking. External methods make use of standalone measurement equipment for localization, whereas image-based methods require image analysis.

TABLE 2. Typical measurement accuracies of active external tracking methods

Tracking method	Measurement volume/range	Position accuracy (RMS)	Orientation accuracy (RMS)
Electromagnetic (Ascension Technologies Inc.)	Sphere, 60 cm radius	0.7-1.4 mm	0.5-1.0°
Optical (Northern Digital Inc.)	Cone, 1.8 × 1.8 m base × 1.5 m height	0.1-0.15 mm	N/A
Mechanical (Immersion Corp.)	Hemisphere, 60 cm radius	0.23-0.43 mm	0.01°

RMS, root mean square; N/A, not applicable.

External Tracking. External tracking systems (as described above in the section on multimodal imaging) report the position and orientation of specially designed tracking objects with respect to a fixed base. By rigidly mounting the tracking objects to surgical instruments, the instruments can be localized within the tracking system coordinate frame. To then place the instrument precisely within the field of view of the imaging system, a second set of tracking objects must be placed on the imaging probe, for the case of ultrasound, and the two coordinate systems coregistered for precise overlay of an indicator or cartoon of the instrument within the operative field (Fig 2). For the case of MRI, the magnet serves as the reference frame for the image coordinates and the instrument tracking positional information must be registered to the image coordinates. Currently available tracking systems use electromagnetic, optical, or mechanical localization techniques. Electromagnetic tracking systems (eg, Ascension Technologies Inc.; Northern Digital Inc.) transmit oscillating magnetic fields from coils contained in a stationary base, and measure the fields at mobile localization devices containing smaller receiving coils. Field strengths at the receiving coils are used to compute the relative position and orientation of the localization device with respect to the stationary base. Although this method does not require line of sight to operate, it is sensitive to ferromagnetic objects, which can cause significant distortion in the operating room environment and are not applicable for MRI.[18] Optical tracking (eg, Northern Digital Inc.; Claron Technology Inc.) uses stereo vision techniques to determine the position of target objects. Multiple tracking targets are combined in fixed patterns for localization.[19] This approach requires minimal obstructions between cameras and targets, so objects inside the body cannot be tracked optically. Mechanical tracking (Immersion Corp.) uses articulated arms attached to a stylus. Measurement of the arm's joint rotations determines endpoint position and orientation of the stylus tip. In this case, the tip is equivalent to the tracking objects used in electromagnetic and optical

TABLE 3. Accuracies of image-based tracking methods

Tracking method	Position accuracy (RMS)	Orientation accuracy (RMS)
Hough Transform (Ding and Fenster, 2003)[21]	N/A	1°
Principal Component Analysis (Draper et al, 2000)[23]	N/A	1°
Passive Tracking Object (Susil et al, 1999)[34]	0.38 mm in CT imaging	0.32° in CT imaging
Active Tracking Object (Merdes and Wolf, 2001)[40]	0.22 to 0.47 mm	N/A

RMS, root mean square; N/A, not applicable; CT, computed tomography.

systems. Measurement accuracies of these three systems are summarized in Table 2.

Image-Based Instrument Tracking. Image-based instrument tracking makes use of the instrument shape as it appears in the image. This approach has three advantages over other tracking methods. First, no calibration procedure is required. Second, if the instrument is tracked with respect to an external base frame then the image can also be tracked with respect to this base frame as well. Third, this method is scalable. Although active external tracking systems must be redesigned for various task length scales, both in form factor and in precision, the principles of image-based tracking scale in proportion to the size of the image. For example, ultrasound scanners are produced in a wide variety of fields of view and resolutions, so imaging of arbitrarily sized tasks is easily achieved.

To accomplish image-based tracking, instrument features must be identified or segmented from the image. Many image segmentation techniques are available for ultrasound images.[20] Techniques for instrument segmentation and tracking can be separated into three categories: pure image analysis, instrument modification for visibility, and attachment of tracking objects. Table 3 summarizes the accuracies of these various methods.

Image analysis methods, which have been employed for instrument segmentation, are the Hough transform, principal component analysis, and texture matching. The Hough or Radon transform[21,22] and principal component analysis[23,24] are geometric segmentation techniques used for finding the instrument shaft within ultrasound images. These algorithms take advantage of the instrument geometry, a straight line, which is different than any other structure in the field of view, to identify and track the instrument. Texture-based segmentation uses statistical metrics of pixel intensity in small neighborhoods to segment instruments from
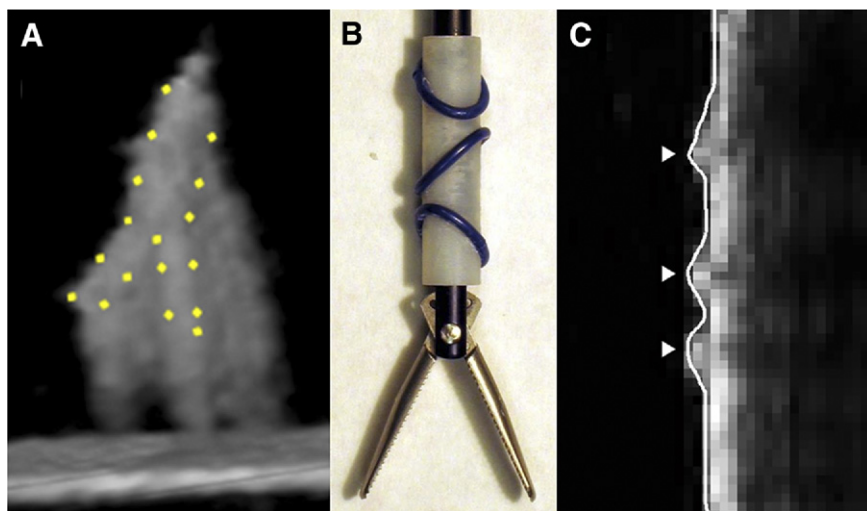
FIG 3. Ultrasound imaging of forceps. (A) Three-dimensional image of forceps with actual location indicated by dots. (B) Fiducial marker on shaft of forceps. (C) Ultrasound image showing detected locations of marker ridges (arrowheads). (Color version of figure is available online.)

tissue.[25] In this method, the texture properties of both instruments and tissue are initially established using a set of presegmented training images. Additional processing is required to identify specific instrument features. Although instrument segmentation techniques can identify some instrument features, such as the instrument shaft, shape distortions can make more complex features such as the jaws of a grasper or scissors impossible to identify (Fig 3).

Instrument modification techniques redesign the instruments to make them easier to segment. For ultrasound imaging, surface treatments can be used to make instruments appear brighter by increasing their scattering. These include scoring or roughening the instrument surface,[26,27] adding patterned contours such as screw threads,[28] and applying various surface coatings containing bubbles.[29,30] More complex surface treatments can apply sophisticated acoustic modeling to further improve instrument appearance. An example is the quadratic residue diffuser,[31] which consists of a pattern of wells drilled into the surface; depths are specified such that sound is evenly scattered from the surface. Additional techniques make use of alternate imaging modes. For example, Doppler imaging can be used to highlight vibrating needles.[32,33]

Finally, tracking shapes can be rigidly attached to instruments and designed to be easily segmented from the image and convey their own local coordinate

frame despite instrument appearance ([Fig 3](#)). This has been implemented using CT[34,35] and in ultrasound.[36] Active tracking shapes, which interface directly with the imaging system, can also be used. This approach involves attaching miniature transducer crystals to the instrument for ultrasound imaging, which can be localized within the image field of view.[37-40] For MR imaging a radiofrequency coil can be attached to the tip of the instrument or catheter for localization within the image.

Image Display

To perform procedures inside the body without direct visual access to the target structures, a display system is required to convey the critical anatomic and functional information to the surgeon in a timely manner and with minimum interruption of the workflow. Depending on the task to be performed and the nature of the target, the challenge is to determine what information is needed, how it should be best displayed, and at what speed the information must be updated (image frame rate). For example, in procedures manipulating structures that are rigid such as bone and cartilage, anatomic information obtained *a priori* can either be displayed for immediate viewing, or it can be registered to the body and displayed as an overlay on a real-time image. Examples of this are overlays of preoperative MRI of the brain overlaid on the image of the skull to plan the location of a craniotomy. For soft tissue or moving structures such as blood vessels, the bladder, or the heart, a more dynamic system is required that provides the surgeon updated information about the position of the target organ. Depending on the speed of motion of either the structure or the surgical tool interacting with the tissue, the frame rate should be appropriately high to permit understanding of the motion characteristics and to prevent inadvertent collision or injury to tissue.

The type of information required for the surgeon to carry out a procedure is task-specific. For some procedures, the surgeon only needs anatomic information to, for example, localize a lesion, such as the site of a tumor in the retroperitoneum. For other procedures, functional information is required such as the area of prolapse in a valve leaflet causing regurgitation. In other procedures, metabolic information is helpful to determine the site of increased activity indicating malignancy and, with still others, capillary density and tissue perfusion is important. Given the wide range of information available, a display system must enable the surgeon to correctly decide the location of the problem and to evaluate the effectiveness of the intervention. Often this information is obtained from different imaging modalities such as CT and MRI along with fluoroscopy or ultrasound. These images then must be coregistered to the anatomic target for
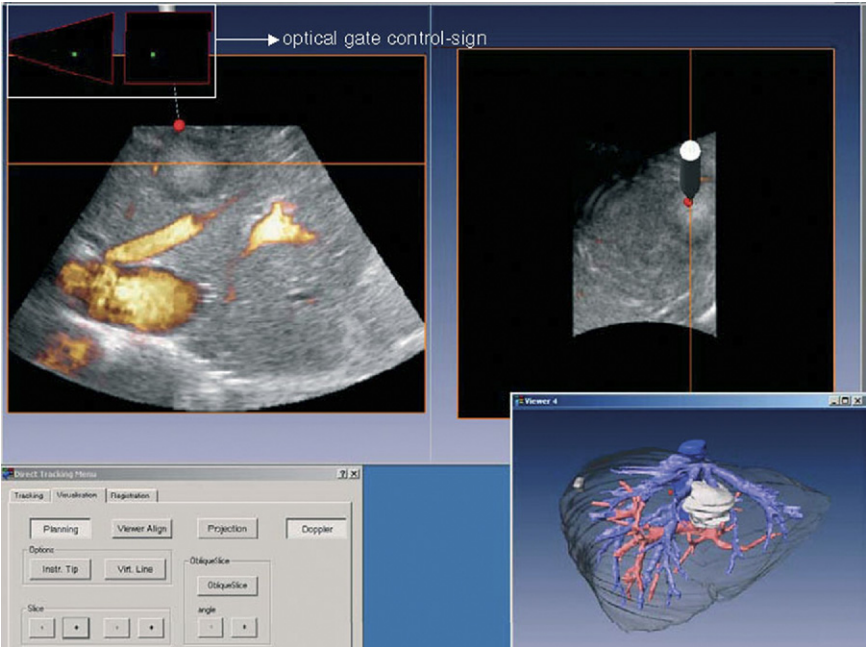
FIG 4. An example of multiple image planes displayed along with the tracked instrument tip to aid navigation and tumor resection. The coronal section plane is visualized in the left display, the transverse section in the right. The levels of the planes are marked with an orange line (in the online version of the figure). The tip of the tracked surgical instrument is visualized with a red spot (in the online version of the figure). By means of this virtual orientation the tumor border is mapped on to the liver and resection can be performed with adequate image-guided distance to the tumor. (Reprinted with permission from Beller et al.[41]) (Color version of figure is available online.)

accurate overlay, and an easily understandable image or set of images must be provided to the surgeon (Fig 4).

When two or more images must be displayed, two options are available. The images can be placed side by side for comparison (Fig 4) or they can be superimposed, making one partially transparent, thus creating an overlay. Furthermore, the option of collecting 2D versus 3D data sets may be available and, for 3D data, there is the question of whether image visualization or rendering should be in 3D or 2D. The fourth dimension, time, also can be critical, particularly when working on rapidly moving structures such as the heart. Thus the science of information display encompasses many areas including display hardware, software development, graphical user interfaces, and human factors to design efficient interfaces, and even display of information about tissue material proper-

ties and surgical tool interaction with tissue through force feedback or haptics. In the next section, we first discuss the various ways of manipulating imaging data, particularly for 3D data sets, and options for displaying the data. This process is often called data visualization. In the second half of this section, we will discuss hardware displays available for the operating room environment.

Visualization of Imaging Data

Visualization of imaging data is the process of transforming image data into a graphic representation that is understandable by the clinician and conveys important information required to assess structure and function. This field has evolved dramatically as the complexity of imaging technology has increased and as 3D and 4D information about tissue and organs has become readily available. Some forms of imaging data require very little if any manipulation to represent the information to the clinician whereas other modes of imaging require complex systems. Projection imaging with x-ray technology is the most commonly used method of imaging and displaying internal structures. Although much information can be conveyed by simply displaying the projection image itself, such as a chest x-ray, positional information and 3D volume information can only be inferred, usually by obtaining images in more than one projection plane (PA and lateral chest radiograph, for example). Fluoroscopy and rapidly updated images displayed on a screen provide time-sensitive information regarding tissue location and motion and instrument position and orientation, but this type of display still leaves it up to the user to mentally integrate the 2D images into a mental 3D image. Although for diagnostic procedures, this information is often adequate, to guide interventional procedures, often more complete visualization of 3D data is required or desirable.

Display of 3D data (CT, MRI, ultrasound). Visualization of 3D imaging data is more complex since there are different options, each one having distinct advantages and different requirements as far as technology used to generate the display. Adding to this complexity is that the human visual system is accustomed to perceive surfaces at various distances but not true volumetric information. For example, perception of one object passing behind another is based on mental models of object shape and inference about how objects move and not on direct volumetric perception. Display of volumetric imaging becomes a matter of providing volumetric information to the user through this limitation.

2D Slices of the 3D Volume. Generating 2D planes that cut through the volumetric (3D) data is a common technique for conveying accurate

spatial information of the area of interest. The planes are generated by creating a linear grid across the 3D volume that can be oriented in any way the operator desires and independent of the orientation by which the 3D data were obtained. Typically these planes are oriented in three directions (axial, sagittal, and coronal views, since they are the most easily understood by clinicians). The perspective, however, can be changed arbitrarily and, in some systems, the cut planes can be made to intersect at the area of interest or even track the area of interest over time, providing the surgeon with quasi-4D information. Although the information provided is limited to the data contained in the specific planes, its main advantage is speed since only limited data are processed and displayed. For computationally intensive systems, such as real-time MRI, current systems are only able to use 2- or 3-slice displays. As computing speeds and software volume rendering algorithms improve, these types of displays will likely become obsolete. Multislice displays can also be integrated with other information obtained from different imaging studies to provide more information about the area of interest.

Three-dimensional data sets require complex algorithms to transform or render the image data into graphical displays. The two options for rendering are surface and volume rendering. Each method of rendering provides different information depending on the needs of the clinician. If the information required is an image of the surface of an organ or structure without the need to know what is below the surface, then surface rendering is ideal since it provides a graphical depiction of the surface details in 3D. If information is required of the surrounding structures, either below the surface or the relationship to nearby structures, then direct volume rendering is required (Fig 4).

Volume Rendering. Volume rendering is the process of assigning each voxel within a 3D volume of image data an opacity or color using a transfer function that is based on the value of each voxel as measured by the imaging detector. The latter is dependent on the imaging technique, since the interaction between the energy used to obtain the image (x-ray, ultrasound, MRI, etc.) and the tissue is different for each imaging technique. Once the process of assigning opacity and/or color to the entire block of data is accomplished, then a decision must be made about the perspective that will be displayed since the data volume can be rotated and displayed in an arbitrary manner. This is sometimes termed the "camera" perspective, since the data are displayed as the view from a camera that is aimed at the object to be displayed. Direct volume rendering can also be described as looking into a fog of varying density. Low density regions will be more transparent than high density regions.
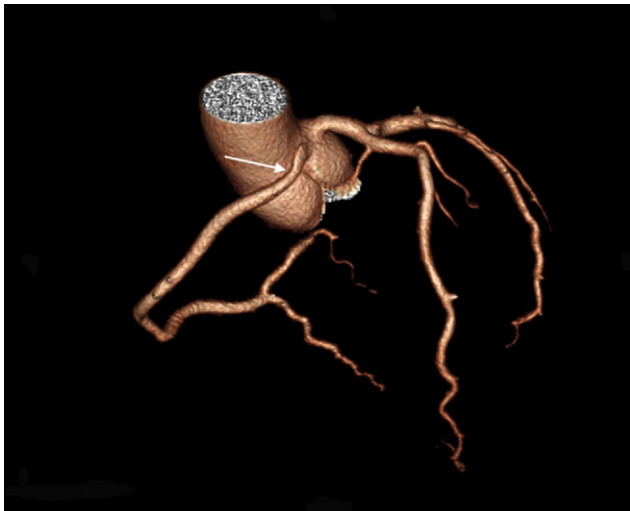
FIG 5. Surface rendered image of aortic root and coronary arteries obtained from a CT scan. Note the intramural course of the right coronary artery (arrow). (Color version of figure is available online.)

The volumetric data are mapped to this density through a transfer function that is application-dependent. An important issue in generating a transfer function is how best to accentuate the boundaries of different tissues. A simple example of this difficulty can be shown by imagining three semitransparent objects in a volume: one blue, one yellow, and one green, representing different types of tissue. When the yellow and blue objects are aligned it will appear that there are two green objects within the volume instead of one. Transfer functions can also be designed to mimic surface rendering by having opacity fall off sharply when the data's value is not similar to that of the structure of interest.

Surface Rendering. Since the human visual system is adept at working with surfaces, a straightforward way to visualize volumetric data is as surfaces. Surface rendering is done by isolating the part of a volume of interest and rendering its surface opaque while making the remainder of the volume transparent. This method works well in instances where there is a clear delineation between the structure of interest and the surrounding tissue. Examples include rendering of bones surrounded by muscles, vessels surrounded by brain matter, or cardiac structures connected to the blood pool. Critical to effective surface rendering is finding a method to clearly delineate the boundary between the structure of interest and the surrounding materials (also known as segmentation). Once the boundary has been clearly identified, it can be displayed by rendering it as a

polygonal mesh. This mesh can be displayed either as a flat 2D on a screen or as a "three dimensional" object using a stereoscopic display.

The most common technique for extracting the boundaries from volume data is the iso-surface. This is a rendering of the volume where all of the data points of a certain value are rendered opaque and the remainder is transparent thus giving the surface where the data have the same or *iso*-value. This is often done in combination with thresholding, where all data values within a range are assigned a value (often 1) and all data outside that range are assigned a different value (often 0). The iso-surface of the data equal to 1 is then displayed (Fig 5).

More powerful image analysis methods can also be used to provide the segmentation for iso-surface generation. For example, statistical analysis can be used to provide cleaner surfaces than simply using the data values directly. A problem with ultrasound images is that intensity values for the same material can vary within an image (ie, tissue can look different in different regions of the image). There are techniques to regularize the appearance of tissue and blood seen in ultrasound images to give a better surface than simple thresholding. These techniques apply statistical methods to determine the probability of a pixel being one form of material versus another (eg, blood vs. tissue). The probability is calculated given the pixel intensity, the normal distributions describing each class of material, given the mean and standard deviation.

Visualization of 3D data. A way to categorize 3D visualization is into stereoscopic and nonstereoscopic. The human visual system uses several cues to determine the placement of objects in space. The most well known of these is binocular (stereoscopic) vision. Binocular vision uses slight differences in views from each eye to determine an object's depth. But this is not the only way the human visual system determines depth and shape. Alternate depth and shape cues include the relative sizes of known objects, texture, lighting/shading/shadows, occlusion, perspective, and motion. There are several nonstereoscopic techniques that rely on these cues to present either volume- or surface-rendered 3D data sets. Nonstereoscopic methods have the advantage that they require no complex stereo display hardware and can be implemented as software solutions. These methods rely on the innate ability of the observer to reconstruct depth information from various cues. These depth and shape cues include relative sizes of known objects, texture, lighting/shading/shadows, occlusion, perspective, and motion.

Perspective projection is also an effective method to show depth, but is most useful either with motion of the objects or with parallel lines within the volume that warp under perspective projection. Motion can be used to

add depth information through user interactions or simple periodic perturbation to the viewpoint. Rotating a view back and forth can be used to show perspective and occlusion depth cues.

Display Hardware—True 3D Displays

Augmented Reality. By providing a display hardware that is semitransparent and registered to the outside environment, data visualization can be used to augment reality. Augmented reality is possible with both stereo and monocular displays. This is done by integrating a real world view (direct or video) with a rendered view. Augmented reality systems have the advantage of providing the surgeon with a real-time view of the actual operative field at the same time as presenting information from other sources either interoperatively (ie, ultrasound) or preoperative CT/MR. The image display can be a simple retro-reflector with a beam-splitter on which the live image data can be displayed, but through which the surgeon can view the operative field. This system, sometimes called suspended image, has been used in neurosurgery for planning a craniotomy; systems exist for laparoscopic surgery as well. Head-mounted displays have also been used for this purpose since they maintain the visual and motor axis of the surgeon. Image resolution and brightness have been a drawback of these systems.

Stereo Display—Software and Hardware. To capitalize on the binocular vision of the human visual system (HVS), special hardware must be used. Although the exact methods vary, all of these systems work by providing a different view of the volume to each of the user's eyes. In addition to the added complexity of the display hardware, stereo-display systems by definition are required to render each view twice. This can be computationally expensive depending on the rendering method used.

There are several stereoscopic displays available for desktop use. These systems can be separated into shutter type and direct display. A shutter-type system relies on the user wearing glasses that will periodically occlude the view of the display device while it is showing the view for the other eye. The advantage of these systems is that they are inexpensive, multiple users can use them at the same time, and they have good resolution (that of the screen being used). The limitations are: limited view as objects are shown with depth but only from one direction so the user cannot "walk around" the object (this is not an issue for a seated user); the inconvenience of having to wear glasses; and the flicker rate, which can be perceived by some users. By their nature, shutter-type devices have half the frame rate of a nonshuttered equivalent display.

Direct stereoscopic displays send separate images to each eye in an

uninterrupted fashion. Traditionally these displays have been either head-mounded displays or dual cathode ray tube (CRT) displays (ie, Da Vinci console; Intuitive Surgical, Inc., Sunnyvale, CA). Consoles offer the greatest resolution since arbitrary CRTs can be used, but severely limit mobility. Head-mounted displays are less limiting but still encumber the user to some extent. The larger disadvantage of head-mounted displays is that they have much lower resolution than console or shutter-type devices. Advances in eye tracking have made it possible to use flat panel displays that use a Fresnel lens to send alternating lines to each eye. The weakness of these displays is that their horizontal resolution is halved compared with other displays. In addition, these displays are limited to one viewer at a relatively close distance.

User Interfaces

User interfaces can be classified broadly into graphical and physical. Although graphical interfaces provide the most flexibility for the system designers, there are several guidance tasks and output types (ie, force feedback) that require specific physical devices as part of the interface. A critical idea in user interface design is "the task-centered user interface design." It is easy to see that image-guided interventions will require different displays than diagnostic displays even though some of the data used in both cases will be the same. In diagnostic display the image quality is more important that the speed at which it is displayed. By contrast during an intervention, timeliness of the presented data can be critical. In addition, a surgeon may not be able to conduct a procedure encumbered by various display devices.

When displays are designed for intervention attention must be paid to efficiency and the comfort of the interface. The discipline of human factors studies the basic performance characteristics of the human body. These studies can be used to provide the best interface under a specific task description. For example, minimum response time is on the order of 150 ms before a user can voluntarily respond to a stimulus. During that time the user will continue motion even if given feedback to stop. During this prevoluntary time, force feedback can be used to resist the direction of motion and can be used to limit incursions into sensitive areas. This is useful in the case of a tele-manipulation, such as with a surgical robot; here the presence of force feedback in the interface can be used to better control motions.

Another important part of the visual interface design is how to provide minimal effort view control to the surgeon. Simply using open verbal communication between the surgeon and an assistant operating the

interface is insufficient for manipulation/navigation. This is because it is difficult to express the desired six degree of freedom positions and orientations verbally.

As interventions move from direct visualization to image-guided, the surgical instruments also change. The interventionalist no longer has direct tactile feedback through the tools. The success of laparoscopic procedures has shown, however, that loss of some tactile information is acceptable. Studies have also shown that visualization can provide force information through sensory substitution. That is to say, if a deformation is perceivable in the visualization, the surgeon's experience can provide an intuition as to the amount of force being applied. In cases where visualization provides insufficient cues, then haptics or force feedback can provide applied force information. The disadvantage of haptic devices, however, is their mechanical complexity.

Complex Displays

To facilitate navigation within large cavities with multiple potential obstacles, such as within the abdomen, complex displays have been constructed that provide navigational aids. An example of what can be done is the concept described as "Anatomical Roadmaps for Diagnostic Navigation in the Abdomen."[42] The goal of this project was to overcome the problem of spatial disorientation when navigating an ultrasound probe over organs and structures in the abdomen. Their solution utilizes a multimodal display combining 3D surface renderings of anatomy, obtained from preoperative imaging, with intraoperative ultrasound. The anatomical models can be constructed from contrast-enhanced CT scans of the patient taken before surgery. The CT volume is first segmented and the models of arterial and skeletal anatomy are generated. Images are then edited to ensure easy separation of different tissue structures. Structures are labeled automatically using a region-identification algorithm. Finally, 3D models of each tissue structure are constructed. The graphical display contains the anatomic model from the segmented CT image, the view from the laparoscope camera, and the ultrasound image obtained from the ultrasound probe inside the abdomen. The position of the tracked ultrasound probe is displayed as an overlay on the anatomic image generated by the CT scan (Fig 6).

Navigation to small-scale structures of interest is performed by mentally recalling their anatomical location with respect to the major vasculature and then using the network of vasculature as a roadmap to navigate the 2D ultrasound probe to the desired location (Fig 6). When this system was evaluated in laboratory studies, although surgeons required more time to
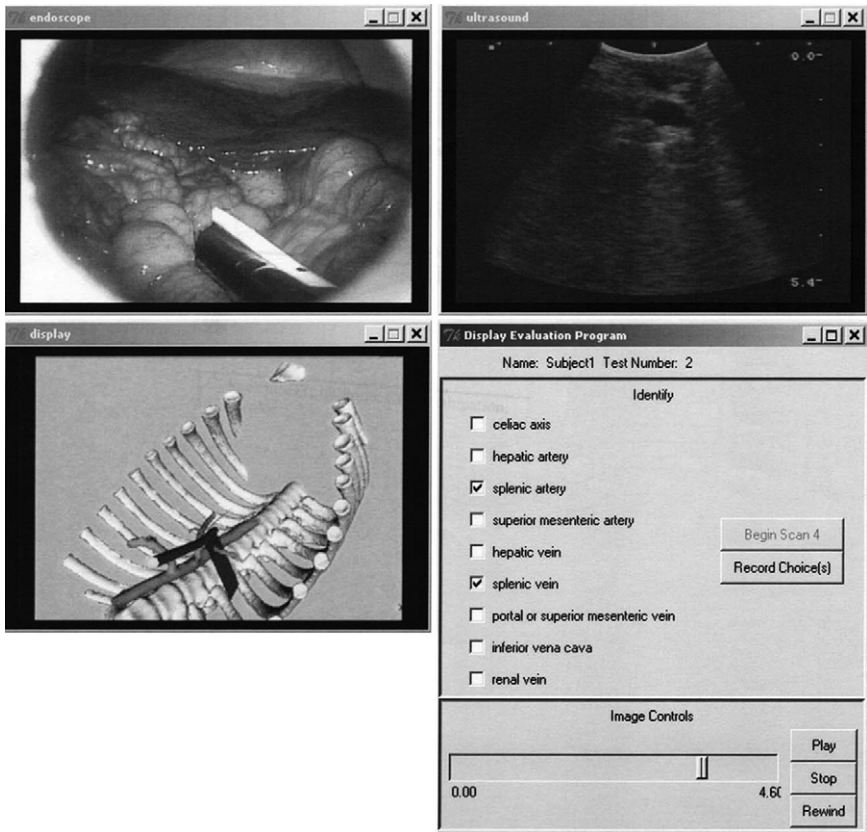
FIG 6. Screen shot of testing software showing laparoscopic view including view of ultrasound probe (upper left), ultrasound image from intra-abdominal probe (upper right), segmented CT image showing probe position and field of view (lower left), and selection panel for segmented CT image. (Reprinted with permission from Ellsmere et al.[42])

analyze the test clips when the graphical display was shown, this extra time resulted in greater image familiarity. It is hypothesized that the time saved in gross navigation with the graphical display offsets the increased effort to understand an extra source of information. An interesting result of this work is that navigation can be performed well to structures that are much smaller than the 3- to 5-mm error in image registration. The registration error need only be small enough to uniquely identify which branches of the vascular roadmap are currently in the ultrasound image. From these large structures, the surgeon is sufficiently spatially oriented to comprehend the much smaller structures shown in the ultrasound images. This result is significant because it shows that more complex,

costly, and time-consuming registration algorithms are not required for accurate navigation.

Specialized Instruments

As for any endoscopic procedure, surgical instruments must be designed to navigate within the confined spaces of the operative field, but one additional factor that is important in image-guided surgery is the ability to detect the location of the surgical tool. Furthermore, for image-guided interventions, the instrument must be made of material that is compatible with the imaging modality. To navigate complex trajectories, instrument flexibility and steerability are very desirable features. Dexterity, or the ability to smoothly manipulate tool position and orientation with as many degrees of freedom as possible, is also an essential feature for most endoscopic procedures. In this section, we discuss current advances in surgical instruments designed for image-guided surgery. Finally, the concept of the surgical tool as an aid to navigation for image-based segmentation and tracking, and as an imaging device or probe, is discussed.

Rigid Versus Flexible Tools

Rigid instruments, in addition to being relatively easy to manufacture, have the added advantage of permitting the surgeon to apply significant force through the instrument shaft to stabilize the instrument against tissue and for tissue displacement, when compared with flexible instruments. Designing dexterous flexible instruments that will have the same endpoint force as their rigid counterparts is a significant challenge. One of the main problems is ensuring the stability at the end of the surgical instrument that is necessary to perform tissue retraction, dissection, and approximation. This problem may be approached by incorporating additional locking mechanisms into the tools. For example, this design has been incorporated into steerable catheters such as those used for mapping and radiofrequency ablation procedures for cardiac arrhythmias.[43] More recently, with increasing interest in natural orifice transluminal endoscopic surgery (NOTES), a similar concept has been used to introduce multifunctional surgical tools through the open channel of a flexible endoscope[44] (Fig 7). Such a design is meant to allow safe passage of the surgical tool into the abdominal cavity toward the anatomic target without damaging neighboring structures and then, by locking the connectors in the instrument, to stabilize the tool tip in the desired orientation and position to perform the procedure. Although this design does provide some stability at the tool tip, it does little to address the need

FIG 7. A "shape-locking" scope shown locked in position. (Reprinted with permission from Spaun et al.[44]) (Color version of figure is available online.)

for application of force at the tool tip to manipulate tissue. Other methods are required for this maneuver. One option is to anchor the instrument to tissue nearby the anatomic target and then to extend the tool tip with the working instruments beyond this point. The stability is then dependent on the rigidity and strength of the nearby tissue and the strength of the instrument-anchoring mechanism.

Steerability and Dexterity

Steerability refers to the ability to control the shape or curvature of the surgical tool as it advances into the body, to navigate around structures. The term has also been used to describe the ability to manipulate the instrument tip through a wide range of motions or degrees of freedom. This latter capability is also called instrument "dexterity" since the goal is to replicate many of the maneuvers that the surgeon makes with their hands in open surgery. Some of the more common options for achieving steerability are with thin cables and pulleys that are controlled by electrical motors or "actuators" and, alternatively, with the use of metal rods connected to the instrument tip by a hinged connection. A potential method for achieving steerability is with the use of superelastic metal alloys, such as nickel-titanium–based instruments. Such instruments take advantage of the unique property of this superelastic alloy that allows it to be bent or deformed significantly without losing its original shape or suffering structural damage. As a simple example, a nickel-titanium
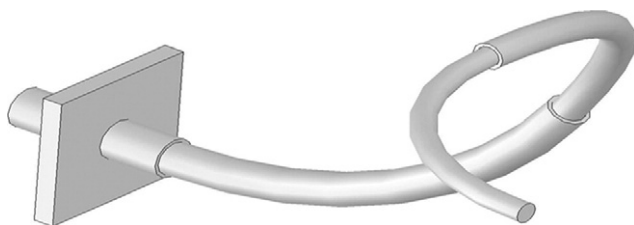
FIG 8. Concentric tube steerable needle. Each tube has a preformed curve and can be rotated and extended to "steer" around objects to a specific target.

instrument with a preformed curve is inserted through a straight rigid tube. Extending the alloy tip generates a variable curvature and, by adding rotation, significant steerability can be achieved.

One of the main disadvantages of the cable-driven locking tip design is the size of the tool. This is due to the size of the mechanism that controls both the steering and the locking. This limitation is important for procedures for which the diameter of the instrument must be very small while still maintaining its steerability. Nickel-titanium instruments, on the other hand, can be made of very small diameter. Furthermore, by having concentric tubes with different curvatures, a steerable instrument that can achieve 6 degrees of freedom can be constructed.[45] By extending and rotating the tubes with respect to each other, the position and orientation of the tip, as well as the shape of the inserted length, can be controlled (Fig 8). Such a system may be used as a delivery device for microinstruments and other repair tools mounted at the tip of the innermost or smallest tube.

Compliant Instruments

Compliant mechanisms are made of monolithic jointless mechanisms that rely on elastic deformation to transmit forces and motion.[46,47] They enable precisely controlled complex movements with more natural hand actions compared with traditional endoscopic tools (Fig 9). The lack of traditional joints offers many design benefits, including the absence of wear debris, pinch points, crevices, and lubrication. An important feature of compliant tools is the built-in return spring action. Force applied in one direction allows operating the device in both, which may complicate the force feedback for an operator. The surgeon feels both the cutting/grasping force and the stiffness of the tool at the same time and has difficulty differentiating them. Despite this limitation, compliant mechanisms may provide a valuable alternative for traditional cable-driven endoscopic tools, especially in limited working space and in the presence of blood.
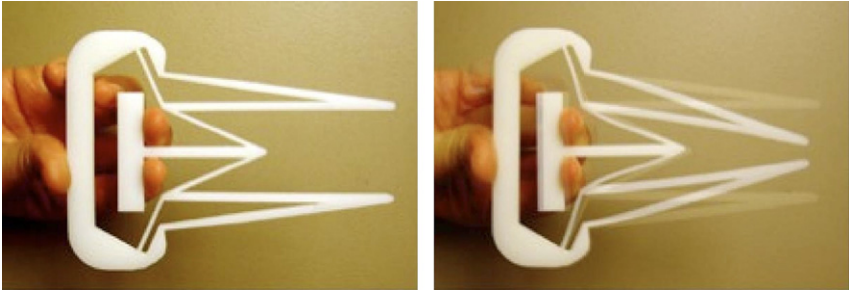
FIG 9. Prototype of the compliant gripper in its inactive mode (left) and gripping mode (right). (Reprinted with permission from Kota et al.[46]) (Color version of figure is available online.)

Motion Compensation Tools

Stabilization of the tool tip is helpful for approaching relatively static structures; however, for interacting with a rapidly moving target such as the beating heart, a different approach must be used. One option is to capture and stabilize the target or organ and, if possible, to immobilize or dampen its movements. Another alternative is to track the movement of the target tissue (image-based tracking) and mimic its motion using a robotically-controlled instrument positioner. The complexity of such a system depends on several factors including the precision of image-based tissue tracking; the motion profile of the tissue, that is how far and how rapidly it is moving in 3D space; and the ability of the instrument positioner to move at the same rate in all 3 directions. Although this is currently achievable only with relatively slow-moving structures, typically noncardiac tissue, there are efforts being made to apply this concept for instrument interaction with specific cardiac structures that have motion primarily in one direction. An example of this is instrument tracking of the mitral valve annulus.[48] Validation experiments using ex vivo models of the moving mitral valve annulus have been reported. With further development, this automated motion tracking system may become a key part of image-guided beating heart surgery, since humans are incapable of performing this complex task without immobilization of the heart.

Imaging-Compatible Tools

The material used to construct surgical instruments often interferes with the imaging modality used. For example, metallic instruments interfere with image acquisition for ultrasound, CT, and, depending on the type of metal, MR imaging. To address this problem several approaches have been studied and developed, most of which are specific to the imaging modality.
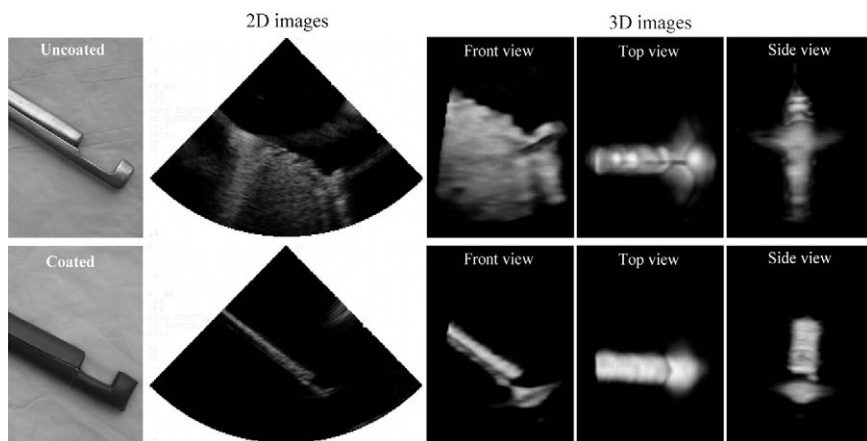
FIG 10. Effect of polyurethane foam coating on suturing device imaging. (Top row) Uncoated. (Bottom row) Coated with polyurethane foam. (Reprinted with permission from Huang et al,[31] ©2007 American Institute of Ultrasound in Medicine.)

Ultrasound-Compatible Instruments. Metallic instruments are problematic to image with ultrasound since sound waves behave very differently in metal objects than they do when propagating through tissue. Their strong reflection and scattering properties generate shadows behind the instrument and, more importantly, produce large artifacts in the image. The artifact can be sufficiently large that it not only obscures the instrument, but also the surrounding structures, making it virtually impossible to determine instrument location and interaction with the target tissue. The extent of the image artifact is dependent on the angle of the ultrasound beam with the surface of the metal instrument. The most common artifacts seen with both 2D and 3D ultrasound are reverberations and side lobe artifacts.[31,49] Reverberations are seen as multiple echoes that can occur when ultrasound is reflected repeatedly inside an object or between two objects, one of which can be the ultrasound transducer. Side lobe artifacts appear due to the presence of ultrasound transducer side lobes, albeit thin in modern machines. Often side lobes appear as specular or diffuse artifactual echoes within the ultrasound beam.[50] To minimize these artifacts, various coating materials on the metallic instrument may be used[17] (Fig 10). Alternatively, the surface of the metal can be modified or roughened to minimize ultrasound reflection.[31] Since the angle of incidence of the ultrasound beam with respect to the instrument is critical, ultrasound transducer position plays a significant role in ultrasound-guided interventions.

MRI-Compatible Instruments. Traditionally, most surgical instruments and devices are made of ferromagnetic materials such as stainless steel. These instruments cannot be used for MRI-guided procedures due to the high magnetic field used for imaging.[51] When placed close to strong magnetic forces without being securely attached to a fixed structure, they can become dangerous projectiles. Other metals, such as aluminum, generate eddy currents, which may cause image artifacts.[52] Eddy currents may also cause excessive heating of the materials, resulting in burns.[53] Ideal MRI-compatible materials should have magnetic susceptibility similar to that of human tissue and be nonconductive. Manufacturability and structural stiffness must also be considered. Thus, a complex combination of plastic, ceramic, fiberglass, carbon fiber, and other composites with a limited number of small metal parts have been used for manufacturing MRI-compatible instruments and devices, including robotic manipulators.[51,54]

Multifunctional Tools

Currently available instruments for minimally invasive surgery are designed as single-function tools and have to be exchanged continually during the procedure.[55] In endoscopic procedures, instrument exchange involves up to 30% of procedure time, significantly adding to total time and constantly disrupting the operator from the actual surgical plan. Several investigators and inventors have proposed various designs of dexterous multifunctional tools for endoscopic procedures with mixed success.[55,56] Although tool multifunctionality has the advantage of having a single tool for various tasks performed inside patient's body, the design concept increases device complexity, particularly at the instrument tip and handle mechanisms. In addition, instrument manufacture and assembly is more time and labor consuming, which increases the price. However, despite these drawbacks such multifunctional tools may become extremely valuable for specific procedures in which instrument navigation is difficult and collision with tissue structures potentially dangerous. An example of such a procedure is intracardiac beating-heart repair where frequent instrument exchange may lead to procedure complications such as blood loss, air embolism, and serious arrhythmias due to instrument collision with tissue.

An additional feature to increase the functionality of the instrument is to incorporate an imaging modality. For example, a multifunctional endoscope with two or more working channels has been developed for gastroenterologic procedures (Fig 7). Since the position of the
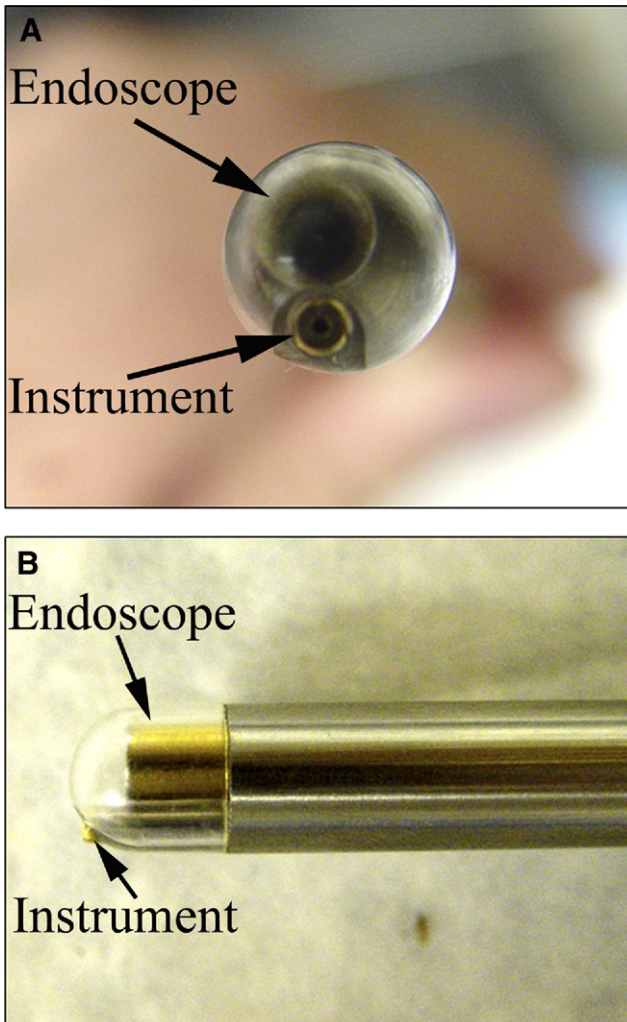
FIG 11. Cardioscopic port. The endoscope (arrow) is inserted into the inner compartment of the port, and the instrument (arrow) is inserted into the working channel. (A) Front view; (B) side view. (Reprinted with permission from Vasilyev et al.[58]) (Color version of figure is available online.)

channels is constant with respect to the image, it is easier to maintain orientation even when the image of the target is upside down.[57] For cardiac applications, there is no commercially available multifunctional instrument. Vasilyev and colleagues, however, have reported development of a "cardioscope" that combines video-assisted optical cardioscopy (VAC) and an instrument channel to access structures

inside the beating heart[58] (Fig 11). In these experiments, epicardial real-time 3D echocardiography is used for navigation and to image the multifunctional tool and cardiac target. The optical channel contained in the instrument is used to image the cardiac structure by pressing the scope against the tissue, displacing the blood, and permitting optical imaging of the heart surface.

Conclusion

Image-guided surgery is the process of adapting sophisticated imaging technology to carry out surgical interventions inside the body that would normally require open procedures with direct inspection of the target area. For surgeons to take full use of available technology, they must first understand the strengths and limitations of the various imaging modalities. They must also be willing to develop procedures that incorporate the imaging information, and utilize specialized instruments that are optimized for these procedures. As the technology advances and navigational aids improve, surgeons will be able to undertake ever more complex repairs not only inside body cavities but also inside solid organs and even rapidly moving organs such as the heart.

Appendix

Glossary of Terms

- *Display*—a system or method of conveying critical anatomic and functional information to the surgeon in a timely manner during a procedure.
- *Fiducials*—objects or structures that can be easily visualized and tracked by a particular imaging modality. Usually used for image registration.
- *Fluoroscopy*—real-time updating of x-ray–generated images from flat-panel semiconductor detectors.
- *Gating*—signal (eg, image) acquisition using a detectable triggering event.
- *Image registration*—refers to the aligning of coordinate frames of two or more image sets from different imaging modalities to produce a composite or multimodal image.
- *Image contrast*—ability to differentiate tissue boundaries; refers to the difference in signal among differing regions.

- *Instrument dexterity*—the ability to smoothly and easily position and orient the tip of a surgical instrument.
- *Instrument navigation*—refers to the ability to advance a surgical tool or catheter within the body, toward a target structure with a controlled trajectory, usually aimed at minimizing collateral injury.
- *Instrument tracking*—a method for localizing a surgical instrument with respect the field of view or the operative target.
- *Real-time imaging*—image display refreshes at 25 to 30 frames per second or faster.
- *Segmentation*—the process of identifying an object or structure. Image segmentation refers to separating regions of a digital image into areas with identical or similar characteristics such as anatomic structures, cavities, or surgical instruments.
- *Spatial resolution*—the ability to detect point targets as distinct; depends on the imaging path to the target. Can refer to 1-, 2-, or 3-dimensional imaging.
- *Tomography*—imaging by sections ($\tau o\mu o\sigma$ is Greek for "section"). Imaging can be made "sharper" by focusing the x-ray beam at a level corresponding to a depth of interest.
- *Tracking system*—a method for determining the precise location (usually in 3D) of a point target or fiducial within the operative field. Tracking systems can use electromagnetic waves, visible light, or a mechanical linkage. Image-based tracking refers to the use of image processing to identify and track an object within the imaged field of view.
- *Visualization of imaging data*—the process of transforming image data into a graphic representation that is understandable by the clinician and conveys important information required to assess structure and function.

REFERENCES

1. Maintz JB, Viergever MA. A survey of medical image registration. Med Image Anal 1998;2:1-36.
2. Zhang XZ, Lu Y. CT-guided transcutaneous fine-needle aspiration biopsy. A report of 350 cases. Chin Med J (Engl) 1990;103:599-603.
3. Serra L, Nowinski WL, Poston T, Hern N, Meng LC, Guan CG, et al. The brain bench: virtual tools for stereotactic frame neurosurgery. Med Image Anal 1997;1:317-29.
4. Schreiner S, Galloway RL Jr, Lewis JT, Bass WA, Muratore DM. An ultrasonic approach to localization of fiducial markers for interactive, image-guided neurosurgery–part II: Implementation and automation. IEEE Trans Biomed Eng 1998;45:631-41.
5. Somer EJ, Mardsen PK, Benatar NA, Goodey J, O'Doherty MJ, Smith MA.

PET-MR image fusion in soft tissue sarcoma: accuracy, reliability and practicality of interactive point-based and automated mutual information techniques. Eur J Nucl Med Mol Imaging 2003;30:54-62.

6. Rohling R, Gee A, Berman L. Three-dimensional spatial compounding of ultrasound images. Med Image Anal 1997;1:177-93.

7. Rohling R, Gee A, Berman L. A comparison of freehand three-dimensional ultrasound reconstruction techniques. Med Image Anal 1999;3:339-59.

8. Carr JC, Fright RW, Gee AH, Prager RW, Dalton KJ. 3D shape reconstruction using volume intersection techniques. Sixth International Conference on Computer Vision (IEEE Cat. No. 98CH36271). New Delhi and Bombay, India: Narosa Publishing House; 1998.

9. Herring JL, Dawant BM, Maurer CR Jr, Muratore DM, Galloway RL, Fitzpatrick JM, et al. Surface-based registration of CT images to physical space for image-guided surgery of the spine: a sensitivity study. IEEE Trans Med Imaging 1998;17:743-52.

10. Muratore DM, Russ JH, Dawant BM, Galloway RL Jr. Three-dimensional image registration of phantom vertebrae for image-guided surgery: a preliminary study. Comput Aided Surg 2002;7:342-52.

11. Penney TB, Mecklinger A, Nessler D. Repetition related ERP effects in a visual object target detection task. Brain Res Cogn Brain Res 2001;10:239-50.

12. Vannier MW, Haller JW. Navigation in diagnosis and therapy. Eur J Radiol 1999;31:132-40.

13. Peters TM. Image-guidance for surgical procedures. Phys Med Biol 2006;51:R505-40.

14. Yaniv Z, Cleary K. Fluoroscopy based accuracy assessment of electromagnetic tracking. Proc SPIE 2006;6141, 61410L.

15. Wells PNT. Advances in ultrasound: from microscanning to telerobotics. Br J Radiol 2000;73:1138-47.

16. Mortensen MB, Fristrup C, Ainsworth A, Pless T, Larsen M, Nielsen H, Hovendal C. Laparoscopic ultrasound-guided biopsy in upper gastrointestinal tract cancer patients. Surg Endosc 2009 April 9 (Epub ahead of print).

17. Huang J, Dupont PE, Undurti A, Triedman JK, Cleveland RO. Producing diffuse ultrasound reflections from medical instruments using a quadratic residue diffuser. Ultrasound Med Biol 2006;32:721-7.

18. Birkfellner W, Watzinger F, Wanschitz F, Enislidis G, Kollman C, Rafolt D, et al. Systematic distortions in magnetic position digitizers. Med Phys 1998;25:2242-8.

19. West JB, Maurer CR Jr. Designing optically tracked instruments for image-guided surgery. IEEE Trans Med Imaging 2004;23:533-45.

20. Noble JA, Boukerroui D. Ultrasound image segmentation: a survey. IEEE Trans Med Imaging 2006;25:987-1010.

21. Ding M, Fenster A. A real-time biopsy needle segmentation technique using Hough Transform. Med Phys 2003;30:2222-33.

22. Novotny PM, Stoll JA, Vasilyev NV, del Nido PJ, Dupont PE, Zickler TE, et al. GPU based real-time instrument tracking with three-dimensional ultrasound. Med Image Anal 2007;11:458-64.

23. Draper KJ, Blake CC, Gowman L, Downey DB, Fenster A. An algorithm for automatic needle localization in ultrasound-guided breast biopsies. Med Phys 2000;27:1971-9.

24. Novotny PM, Cannon JW, Howe RD. Tool localization in 3D ultrasound images. Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2003, Montreal, Canada. Montreal: Springer, 2004:969-70.

25. Linguraru MG, Vasilyev NV, del Nido PJ, Howe RD. Atrial septal defect tracking in 3D cardiac ultrasound. Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv 2006;9(Pt 1):596-603.

26. Hopkins RE, Bradley M. In-vitro visualization of biopsy needles with ultrasound: a comparative study of standard and echogenic needles using an ultrasound phantom. Clin Radiol 2001;56:499-502.

27. Nichols K, Wright LB, Spencer T, Culp WC. Changes in ultrasonographic echogenicity and visibility of needles with changes in angles of insonation. J Vasc Interv Radiol 2003;14:1553-7.

28. Reading CC, Charboneau JW, James EM, Hurt MR. Sonographically guided percutaneous biopsy of small (3 cm or less) masses. AJR Am J Roentgenol 1988;151:189-92.

29. Gottlieb RH, Robinette WB, Rubens DJ, Hartley DF, Fultz PJ, Violante MR. Coating agent permits improved visualization of biopsy needles during sonography. AJR Am J Roentgenol 1998;171:1301-2.

30. Culp WC, McCowan TC, Goertzen TC, Habbe TG, Hummel MM, LeVeen RF, et al. Relative ultrasonographic echogenicity of standard, dimpled, and polymeric-coated needles. J Vasc Interv Radiol 2000;11:351-8.

31. Huang J, Triedman JK, Vasilyev NV, Suematsu Y, Cleveland RO, Dupont PE. Imaging artifacts of medical instruments in ultrasound-guided interventions. J Ultrasound Med 2007;26:1303-22.

32. Jones CD, McGahan JP, Clark KJ. Color Doppler ultrasonographic detection of a vibrating needle system. J Ultrasound Med 1997;16:269-74.

33. Jandzinski DI, Carson N, Davis D, Rubens DJ, Voci SL, Gottlieb RH. Treated needles: do they facilitate sonographically guided biopsies? J Ultrasound Med 2003;22:1233-7.

34. Susil RC, Sobie EA, Tung L. Separation between virtual sources modifies the response of cardiac tissue to field stimulation. J Cardiovasc Electrophysiol 1999;10:715-27.

35. Masamune K, Fichtinger G, Patriciu A, Susil RC, Taylor RH, Kavoussi LR, et al. System for robotically assisted percutaneous procedures with computed tomography guidance. Comput Aided Surg 2001;6:370-83.

36. Stoll J, Dupont P. Passive markers for ultrasound tracking of surgical instruments. Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv 2005;8(Pt 2):41-8.

37. Breyer B, Cikes I. Ultrasonically marked catheter—a method for positive echographic catheter position identification. Med Biol Eng Comput 1984;22:268-71.

38. Winsberg F, Mitty HA, Shapiro RS, Yeh HC. Use of an acoustic transponder for US visualization of biopsy needles. Radiology 1991;180:877-8.

39. Vilkomerson D, Lyons D. A system for ultrasonic beacon-guidance of catheters and other minimally-invasive medical devices. IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control 1997;44:496-504.

40. Merdes CL, Wolf PD. Locating a catheter transducer in a three-dimensional ultrasound imaging field. IEEE Trans Biomed Eng 2001;48:1444-52.

41. Beller S, Hünerbein M, Lange T, Eulenstein S, Gebauer B, Schlag PM. Image-

guided surgery of liver metastases by three-dimensional ultrasound-based optoelectronic navigation. Br J Surg 2007;94:866-75.

42. Ellsmere J, Stoll J, Wells W 3rd, Kikinis R, Vosburgh K, Kane R, et al. A new visualization technique for laparoscopic ultrasonography. Surgery 2004;136:84-92.

43. Calkins H, Hall J, Ellenbogen K, Walcott G, Sherman M, Bowe W, et al. A new system for catheter ablation of atrial fibrillation. Am J Cardiol 1999;83:227D-36D.

44. Spaun GO, Zheng B, Martinec DV, Cassera MA, Dunst CM, Swanstrom LL. Bimanual coordination in natural orifice transluminal endoscopic surgery: comparing the conventional dual-channel endoscope, the R-Scope, and a novel direct-drive system. Gastrointest Endosc 2009 May;69:e39-45.

45. Sears P, Dupont P. Inverse kinematics of concentric tube steerable needles. Proceedings of the IEEE International Conference on Robotics & Automation, Rome, Italy, April 2007.

46. Kota S, Lu KJ, Kreiner K, Trease B, Arenas J, Geiger J. Design and application of compliant mechanisms for surgical tools. J Biomech Eng 2005;127:981-9.

47. Frecker MI, Schadler J, Haluck RS, Culkar K, Dziedzic R. Laparoscopic multifunctional instruments: design and testing of initial prototypes. JSLS 2005;9:105-12.

48. Kettler D, Plowes R, Plowes RD, Novotny PM, Vasilyev NV, del Nido PJ, et al. An active motion compensation instrument for beating heart mitral valve surgery. Proceedings of the IEEE International Conference on Intelligent Robots and Systems–IROS, San Diego, California, October 29-November 2, 2007.

49. Hedrick WR, Peterson CL. Image artifacts in real-time ultrasound. J Diag Med Sonog 1995;11:300-8.

50. Laing FC, Kurtz AB. The importance of ultrasonic side-lobe artifacts. Radiology 1982;145:763-8.

51. Tsekos NV, Khanicheh A, Christoforou E, Mavroidis C. Magnetic resonance-compatible robotic and mechatronics systems for image-guided interventions and rehabilitation: a review study. Annu Rev Biomed Eng 2007;9:351-87.

52. Shellock FG. Radiofrequency energy-induced heating during MR procedures: a review. J Magn Reson Imaging 2000;12:30-6.

53. Dempsey MF, Condon B, Hadley DM. Investigation of the factors responsible for burns during MRI. J Magn Reson Imaging 2001;13:627-31.

54. Elhawary H, Zivanovic A, Davies B, Lampérth M. A review of magnetic resonance imaging compatible manipulators in surgery. Proc Inst Mech Eng [H] 2006;220:413-24.

55. Frecker MI, Powell KM, Haluck R. Design of a multifunctional compliant instrument for minimally invasive surgery. J Biomech Eng 2005;127:990-3.

56. Pietrabissa A, Dario P, Ferrari M, Stefani C, Menciassi A, Moretto C, et al. Grasping and dissecting instrument for hand-assisted laparoscopic surgery: development and early clinical experience. Surg Endosc 2002;16:1332-5.

57. Rattner D, Kalloo A. ASGE/SAGES Working Group on Natural Orifice Transluminal Endoscopic Surger, October 2005. Surg Endosc 2006;20:329-33.

58. Vasilyev NV, Martinez JF, Freudenthal FP, Suematsu Y, Marx GR, del Nido PJ. Three-dimensional echo and videocardioscopy-guided atrial septal defect closure. Ann Thorac Surg 2006;82:1322-6, discussion 1326.