7 Interaction in Augmented Reality Image-Guided Surgery

Simon Drouin, D. L. Collins, and M. Kersten-Oertel

CONTENTS

7.1 INTRODUCTION

The goal of augmented reality (AR) in image-guided interventions is to allow surgeons to access navigation information during a procedure without shifting their attention away from the operative field. To be effective, augmentation should provide the right information at the right time, avoid distracting the surgeons, and provide an unambiguous and perceptually sound representation of the preoperative plans, guidance information, and anatomy under the visible surface of the patient. In this chapter, we argue that to reach these goals, it is necessary to systematically consider user interaction in the design of new surgical AR methods.

User interaction design falls under the domain of human–computer interaction and describes the design of interactions between users and their products. The goal of interaction design is to develop methods that allow users to achieve their objectives in the best way possible. Some of the elements developers need to consider when designing interactions include the visual representations of the data, the hardware device that should be used (e.g., mouse), the physical space in which the user will interact with the system, and the behaviors of the user (i.e., how they perform actions and operate the system). To date, there has been little focus on designing effective and usable interactions in image-guided surgery (IGS), and in particular in AR image-guided surgery, despite an important relationship between AR and interaction.

To illustrate the codependency of user interaction and AR, we classify interactive AR techniques into three categories: *control*, *task simplification*, and *enhancing perception*. Interactions for the purpose of *control,* as the name suggests, are intended to manipulate or command the IGS system itself. This may involve simple tasks, such as loading anatomical volumes or plans; enabling functions such as registration and segmentation; and changing visualization parameters. The purpose of *task simplification* interactions is to simplify and enable tasks to be executed more easily, more naturally, more quickly, and/or more precisely. Examples of this include easier localization of anatomy and more accurate targeting. Lastly, *enhancing perception* interactions refer to the use of interaction to improve perception of the anatomy, surgical plans, and, in general, the data visualized on the IGS system. Examples in this area include interactions that allow a user to view occluded objects and understand spatial relationships between data.

The remainder of this chapter is organized as follows. First, we establish the basis that supports our claim about the importance of considering interaction when designing AR visualization tools. Next, in Section 7.3, we review the literature on AR in IGS that pertains to some form of user interaction. Finally, in Section 7.4, we look at possible areas of focus for the future.

7.2 ON THE NECESSITY TO CONSIDER INTERACTION

AR and interaction complement each other in different ways, depending on the aspect of the system that is considered or the specific task that is addressed. In this section, we examine how the codependency of AR and interaction unfolds in light of our categorization of AR methods into *control* of the system, *task simplification*, and *enhancing perception*.

7.2.1 Control

Interaction may be used simply to manipulate or *control* an IGS system and its parameters, for example, to allow the surgeon to control the type of visualization and the displayed content at a given point in time. Such control by the surgeon is important, as AR has been shown to be a source of distraction during an operation (Dixon et al. 2013). This effect can be mitigated by ensuring that the display always reflects what is needed by the surgeon. Although research in the area of surgical process modeling, such as the studies conducted by Jannin et al. (2003), Lalys and Jannin (2014), and Forestier et al. (2013), may help to automatically show the right information at the right time, automatic systems may never be able to predict the surgeon's needs all the time (Dergachyova et al. 2016). For this reason, it is important to give the surgeon control over the IGS system and specifically the AR view, allowing them to control what, how, and when content is displayed.

One of the promises of AR in IGS is to allow the surgeon to access navigation information without shifting attention from the operating field. In this chapter, we argue that this promise cannot be completely fulfilled unless the surgeon is able to also control the display without looking away.

7.2.2 TASK SIMPLIFICATION

One of the goals of IGS systems is to *simplify tasks* such as positioning and orienting surgical tools relative to a target. This form of interaction, when performed with a conventional IGS system, involves a dissociation of the frames of references of the visual and motor systems of the user (Masia et al. 2009). Such dissociation has received much attention in the field of experimental psychology (Harris 1965) and is known to cause an important degradation in the performance of a subject during different motor tasks, such as aiming and reaching, particularly when frames of references are rotated relative to each other (Abeele and Bock 2001), which is often the case in surgery. In task simplification procedures, interaction with the IGS system is improved by the use of AR, which provides the important service of aligning the frames of reference of the visual and motor system.

7.2.3 ENHANCING PERCEPTION THROUGH USER INTERACTION

One of the most frequently reported problems of current medical AR systems is the degradation of depth perception that occurs when combining real and virtual elements. We argue that interaction can help to compensate for this loss of depth perception.

In his book about human visual perception, Gregory (1977) writes, “Perception is not determined simply by the stimulus patterns; rather it is a dynamic searching for the best interpretation of the available data.” This quote highlights how visual perception is essentially a dynamic process through which different pictorial cues are fused by the brain to build a complete representation of the 3D world that is consistent with all cues (Landy et al. 1995). Some of these cues are dynamic by nature. For example, the well-studied motion parallax depth cue (Buckthought et al. 2014), which is generated by head movement relative to the environment, is one such cue where the brain integrates optical flow on the retina over time. Other cues are usually labelled as “static” because they can provide information based on the stimulus patterns captured at a single time point. However, it is often the case that integration over time produces a more complete description of the environment (Johnston et al. 1994).

In the field of cognitive science, researchers have long studied the interaction between visual and motor systems in the human brain. Wexler and van Boxtel (2005) have reviewed how an observer’s motor actions influence his or her visual perception of the three-dimensional (3D) environment. Several results reported by Wexler support the importance of considering interaction when designing visualization tools. For example, observers can recognize objects more easily from a novel point of view if the change of point of view is the result of self-motion, rather than if the object itself was moved (Simons et al. 2002). This is an important consideration for IGS interaction design, as often it is not the surgeon but another member of the surgical team who interacts with the system, which therefore may cause a disruption in understanding the anatomical data. Harman et al. (1999) showed that "observers who actively rotated novel, three-dimensional objects on a computer screen later showed more efficient visual recognition than observers who passively viewed the exact same sequence of images of these virtual objects." Anticipatory mechanisms in the brain are thought to foresee the sensory consequences of motor actions and thus facilitate recognition (Wexler and van Boxtel 2005). For example, blindfolded subjects who are led to a new location can point to objects in their environment more easily than when they imagine walking the same path (Rieser et al. 1986) or when they are shown the corresponding optic flow (Klatzky et al. 1998). Regarding motion parallax, Wexler shows that the same optical flow can lead to different perceptions of 3D shapes depending on whether the motion is generated by the observer or the object itself. Knill (2005) showed that the brain changes how it integrates visual cues based not only on the information content of the stimuli but also on the task for which the information is used. Marotta et al. (1995) showed that monocular observers generate more head movement during reaching tasks to better utilize retinal motion cues.

It is important to understand how AR visualization differs from natural vision and where interaction may help. Drascic and Milgram (1996) review perceptual issues in AR and argue that in an AR environment, technological limitations remain such that certain depth cues commonly used in natural vision are absent, or worse, contradict each other, leading to a distorted perception of depth for the user. However, they show that as more cues are present and are congruent with each other, accurate depth perception can be restored.

In Bingham and Pagano (1998), the authors argue that the study of *definite distance perception* requires a perception-action approach. The reason is that definite distance perception entails calibration, which is a task-specific action that provides feedback to the visual system about the environment. They show that calibration can eliminate underestimation of distance generated by restriction of the visual field. In Altenhoff et al. (2012), it was reported that the typical underestimation of distances under virtual environment navigation can be compensated by submitting users to a calibration session, in which a reaching task is performed with haptics and visual feedback.

7.3 LITERATURE REVIEW

Surprisingly, interaction has not received a lot of attention from the IGS research community. Few publications address the topic of surgeon interaction with IGS systems and even fewer of interacting with visualized data to enhance perceptibility. In their review of AR visualization for IGS, Kersten-Oertel et al. (2013) note that interaction has been limited to allowing the end-user "to rotate and translate objects, to navigate through the virtual scene, to use cutting planes, to toggle components, turn visibility on and off, and to change the opacity and color of objects." Further, the authors found that the types of hardware that are most often used for manipulating data and interacting with the IGS system are the keyboard and mouse. In the following section, we give examples of how interaction has been explored in the context of IGS (rather than presenting a comprehensive review).

7.3.1 CONTROL

The classical keyboard and mouse interaction paradigm is not suitable for interactions between surgeons and surgical augmented environments (Navab et al. 2007; Bichlmeier et al. 2009). A major constraint for interaction, owing to the operating room (OR) environment, is the sterile boundary; when the surgeon is scrubbed, they would need to break asepsis to use the input devices to manipulate the images and system. Therefore, typically it is not the end-user of the IGS system (i.e., the surgeon) who interacts with the system but rather a technician or member of the surgical team receiving verbal instructions from the surgeon. This type of indirect communication is often slow and prone to errors and misunderstandings resulting from verbal ambiguities (Onceanu and Stewart 2011). To overcome these issues, a number of research groups have begun to study natural user interfaces that can interpret human action without direct contact. In AR IGS systems, voice interfaces, gesture-based interfaces, and tracked surgical tools have been explored to control the IGS system.

7.3.1.1 Voice

Few groups have explored the use of voice-based interaction within the OR. One exception is the work by Sudra et al. (2007), who allowed both speech- and gesture-based interaction within the context of an endoscopic robotic system. Their interaction methods allowed the surgeon to use speech or gestures to switch between visualization methods: for example, to change parameters or to turn annotation information on and off. The difficulties of voice interfaces include the challenges of a noisy OR and that voice may not be suitable for manipulation of continuous parameters (Mewes et al. 2017). However, combining voice-based interfaces with gestures may overcome the limitations and shortcomings of both of these types of interactions.

7.3.1.2 Gesture-Based Interaction

A number of groups have explored the possibility of using gestures to allow the surgeon to interact with the IGS system directly while remaining sterile. In a gesture-based interface, a set of motions or configurations of the hands or body are recognized by the system as commands. Wen et al. (2014) developed a gesture-based system for their AR needle guidance system for tumor ablation; three hand gestures are recognized: rotation, translation and point selection (Figure 7.1). A Kinect depth sensor (Microsoft Corp., Redmond, WA) is then used to recognize gestures that enable manipulation of the 3D view.

Kocev and his colleagues developed an interface method for projector-based AR through which the interface of a traditional navigation system is projected on the sterile draping close to the operating field. A Microsoft Kinect was then used to recognize gestures that enabled manipulation of the 3D view (Kocev et al. 2014).

Numerous other groups have explored the use of gesture-based interaction in traditional IGS. For example, Gratzel et al. (2004) proposed a paradigm to replace the use of a mouse in conventional navigation systems with a computer vision-based gesture recognition system that provides similar functionality, and Kirmizibayrak et al. (2011) compared the use of a gesture-based interaction system

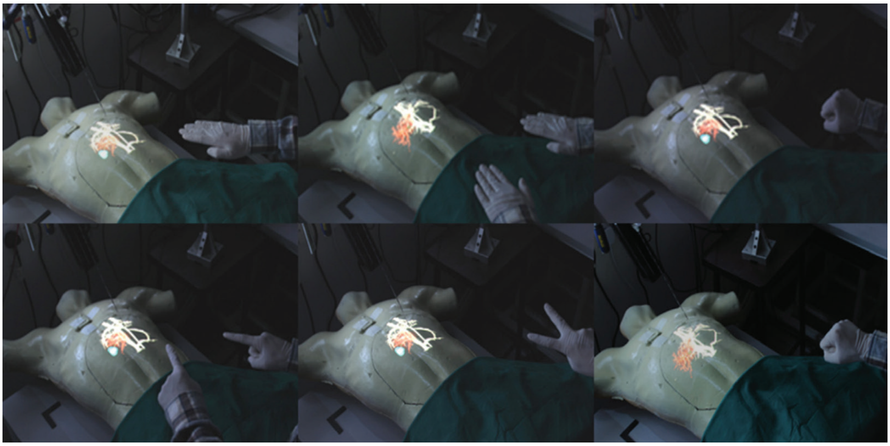

FIGURE 7.1 Gesture-based control of a projected AR navigation system. (From Wen, R. et al., *Comput. Methods Programs Biomed.*, 116, 68–80, 2014.)

with traditional mouse interaction. Although it does not use AR, their system is used for types of interaction that are often found in AR, such as the manipulation of a Magic Lens (Bier et al. 1993), which enables focus and context visualization of multiple volumetric medical datasets. The authors found the Magic Lens gesture interface was faster compared to the mouse but had higher variance. In a similar study with 30 physicians and senior medical students, Wipfli et al. (2016) compared three different interaction modes for image manipulation in the surgical domain: gestures (using Microsoft Kinect), verbal instructions given to a third party, and the mouse. They found that efficiency and user satisfaction were best when using the mouse, followed by gesture-controlled and verbal instructions. Not surprisingly, in their study the mouse outperformed the gesture-based interface, as it was a more familiar tool. For more information about touchless or gesture-based interfaces in surgical environments, the reader is referred to Mewes et al. (2017).

7.3.1.3 Surgical Tool

Another solution to enable the surgeon to interact directly with the IGS system is to make use of the surgical tools within the OR, typically accomplished by providing a way for the surgeon to use the tracked surgical probe. The advantage of using a surgical tool as an interaction solution is that it requires no (or limited) additional hardware in an already crowded OR yet allows the surgeon to have direct control over the IGS system. Examples of this include the work by Salb et al. (2003), who allowed the end-user to interact with a virtual graphical interface through a head-mounted display (HMD) using a tracked Polaris surgical probe. Fischer et al. (2005) use an AR menu that provides selectable menu items and allows interaction using the surgical probe to define points and freely draw shapes in 3D, while Katić et al. (2010) use an AR system that allows the end-user to interact with the system via a 3D pointer with integrated buttons.

In a non-AR IGS system, Onceanu and Stewart (2011) built a joystick-like interaction tool for the OR using a base in which a tracked surgical pointer can be inserted, rotated, and used to click. When comparing their input device with a mouse and verbal communication, the authors found that, although faster than both dictation and the joystick, the mouse was not significantly more accurate. Their results, similar to those of others who have compared novel interaction methods to the mouse and keyboard, suggest that metrics other than accuracy and speed are needed to enable a viable comparison that considers the constraints of the OR.

7.3.2 SIMPLIFYING TASKS

Many surgical procedures require aligning objects or surfaces in 3D, and the solution that is most commonly proposed is to reduce the problem to a two-dimensional (2D) alignment task, which is much simpler to execute. The most common example involves aligning a tool to an axis. Diotte et al. (2015) simplify the interlocking of intramedullary nails by tracking colored markers attached to the drill used to perform the procedure. The drill can then be aligned with overlaid targets and an x-ray image acquired from the same point of view (Figure 7.2a). Herrlich et al. (2017) simplify the needle insertion task by adding a small display to the needle-guiding tool showing a 2D representation of the alignment of the needle with the desired path. Seitel et al. (2016) produce an AR view for percutaneous needle insertion based on a combined color and depth (RGBD) image where the depth information in the image enables registration of the preoperative patient surface and needle plan with the image. A representation of the planned needle path can be superimposed on the image to serve as a guide to accurately position the real needle (Figure 7.2b). State et al. (1996) overlay a 3D representation of an ultrasound (US) image plane with live video of the operating field to guide a needle biopsy. The real part of the image shows the needle insertion point and orientation, while the US image augmentation allows the surgeon to see the target tumor and the

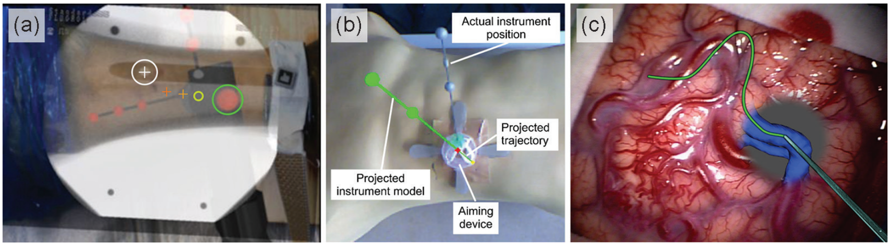

FIGURE 7.2 (a) AR view for interlocking of intramedullary nails: overlay of live video, x-ray, and alignment targets. (From Diotte, B. et al., *IEEE Trans. Med. Imaging*, 34, 487–495, 2015.) (b) AR view for percutaneous needle insertion showing real instrument and alignment target. (From Seitel, A. et al., *Int. J. Comput. Assist. Radiol. Surg.*, 11, 107–117, 2016.) (c) Tracing of virtual blood vessels using the surgical pointer. Trace is then used for registration correction. (From Drouin, S. et al., Interaction-based registration correction for improved augmented reality overlay in neurosurgery, in *Augmented Environments for Computer-Assisted Interventions*, Lecture Notes in Computer Science, Vol. 9365, pp. 21–29, 2015.)

part of the needle that has already been inserted. The fact that both the real needle and the target are displayed in the same 3D context facilitates aiming.

Another very common 3D alignment problem in medical imaging is the registration of the patient with preoperative images and plans. Drouin et al. (2015) propose an AR-based interface to correct the initial patient misregistration in neurosurgical interventions. The interface allows the surgeon to identify corresponding real and virtual anatomical features on the AR view using the tracked pointer typically available in most navigation systems (Figure 7.2c). The registration correction is achieved by automatic alignment of the marked features. An additional advantage of such a method is that it allows the surgeon to visually assess the accuracy of the registration through the AR view. A simpler, but similar, method was presented by Kantelhardt et al. (2015), through which the outlines of pre-segmented brain structures are superimposed on a microscope image of the patient, allowing surgeons to manually modify the initial patient registration to align the outlines with the live video. Cutolo et al. (2015) propose an AR-based interface to help with the manual alignment of non-tracked rigid bodies that need to be accurately positioned on the patient. The rigid bodies can be positioned by aligning virtual points in the AR view with physical landmarks on the object.

AR interaction can also facilitate the localization of a surgical target. Rong Wen et al. (2017) proposed a tablet-based AR system for surgical tool navigation. They evaluated the proposed system and showed that the tablet-based visual guidance system could assist surgeons in locating internal organs, with errors between 1.74 and 2.96 mm, while Shamir et al. (2011) proposed an AR system to explore the safest path for the insertion of straight tools in image-guided keyhole surgery. Their system virtually overlays risk data on a physical model of the patient to facilitate identification of risk-free paths. Shimamura et al. (2013) proposed to use a tracked, handheld display to visualize a slice of the patient's preoperative volume parallel to the display surface. The cutting plane-patient intersection is shown on the skin surface by projecting a laser line from the side of the tablet. As a final example, Bajura et al. (1992) use AR to guide the acquisition of US images. Many other tasks could be simplified by the use of AR, and the body of work on task simplification in conventional image-guidance systems should be thoroughly investigated to find the examples that benefit from adaptation to AR.

7.3.3 Enhancing Perception

As demonstrated above, visual perception in augmented environments can be distorted by the absence of specific depth cues or inconsistencies between various depth cues resulting from the combination of real and virtual graphical objects. Perhaps the most important of these cues is occlusion. Inconsistent occlusion patterns in an image take precedence over other depth cues and lead to an incorrect perception. The most common example is found in AR views in which a real video image of a patient is overlaid on a virtual representation of the underlying anatomy. In this case, a naive alpha-blending between real and virtual components of the image produces the perception that virtual components are floating above the surface of the patient, as illustrated in Figure 7.3a.

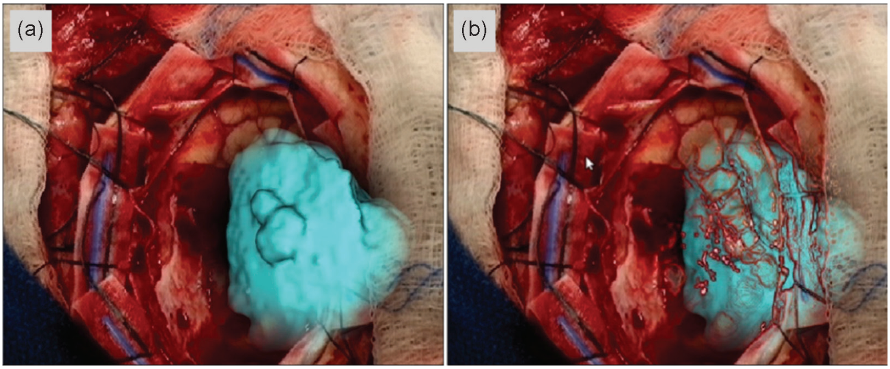

FIGURE 7.3 Example of a medical AR image where (a) the virtual element seems to be floating above the surface for lack of occlusion cues and (b) with occlusion cues restored, the virtual object is perceived to be located behind the surface. (From Kersten-Oertel, M. et al., *Int. J. Comput. Assist. Radiol. Surg.*, 10, 1823–1836, 2015.)

One of the typical approaches to help the visual system solve the occlusion in an AR scene is the virtual window (or Magic Lens) (Bier et al. 1993). Bichlmeier et al. (2007a) cut a virtual window out of the real image by intersecting a 3D model of the patient's skin surface with a user-defined cuboid. The motion of the HMD generates motion parallax and texture accretion/deletion cues to improve perceived relative depth between the surface and the surgical target. The approach was refined in Bichlmeier et al. (2007b) (Figure 7.4b), where the virtual window position is modified according to line of sight of the HMD, and the opacity of the real image inside the window is modulated by the angle between the skin surface and viewing direction as well as the curvature of the skin surface and distance to

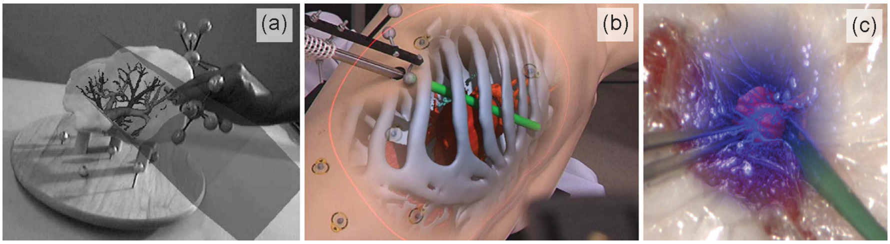

FIGURE 7.4 Different methods using the concept of a virtual window. (a) A tracked virtual cuboid cutout. (From Mendez, E. et al., Interactive context-driven visualization tools for augmented reality, in *IEEE/ACM International Symposium on Mixed and Augmented Reality*, pp. 209–218, IEEE, 2006.) (b) Virtual window based on HMD pose and surface properties. (From Bichlmeier, C. et al., Contextual anatomic mimesis hybrid in-situ visualization method for improving multi-sensory depth perception in medical augmented reality, in *IEEE/ACM International Symposium on Mixed and Augmented Reality*, pp. 1–10, IEEE, 2007; Courtesy of Christoph Bichlmeier.) (c) Force feedback-generated virtual window. (From Gras, G. et al., Visual Force Feedback for Hand-Held Microsurgical Instruments, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Vol. 9349, pp. 480–487, 2015.)

the center of the window. Mendez et al. (2006) propose controlling the pose of a virtual 3D surface using an optical tracker tool. The intersection of the 3D surface with the presegmented organ surfaces is used to create a window where internal organ structures, such as blood vessels, are revealed (c.f., Figure 7.4a). Kalkofen et al. (2009) allow users to manipulate the position of a circular virtual lens, where the virtual window is not simply an area where the real image is semitransparent but instead allows for the application of programmable compositing schemes for the different elements in the scene. Different compositing inside the region covered by the virtual window reveals structures of interest while maintaining contextual information. Gras et al. (2015) (Figure 7.4c) proposed to use a tracked surgical dissector equipped with a force-torque sensor to interact with the AR view. Its location and the measurement of the force applied to the tissues are used to modulate the size of a virtual window, and the *pq* space method of Lerotic et al. (2007) is used to produce realistic occlusion patterns inside the virtual window. Furthermore, Gras and colleagues used force feedback information to distort a virtual model of the organs below the surface, providing a motion cue that improves depth perception even more.

Researchers address the concept of a virtual window in different ways. While Gras et al. (2015), Mendez et al. (2006), and Kalkofen et al. (2009) rely on the user to manipulate the location of the window, Bichlmeier et al. (2007b) use the motion of an HMD, and Bichlmeier et al. (2007a) use a static virtual window and rely entirely on the relative motion of an HMD to create the depth enhancing effect.

Even when employing a concept such as the virtual window, AR may still suffer from occlusion inconsistencies. For example, if an object such as a scalpel is manipulated above the virtual window, the part of the object intersecting the window will be occluded. Kutter et al. (2008) partially solve the problem by tracking the surgeon's hand in the live video stream to produce an occlusion mask for the virtual content of an AR view. Pauly et al. (2015) propose a similar but more general approach that can track arbitrary objects that occlude the operative field from an RGBD camera data using a random forest approach.

One interesting strategy to enhance perception in AR is to create bridges between real and virtual parts of an image to mitigate the influence of depth inconsistencies on perception. Choi et al. (2015) dynamically trace a line in the AR view connecting a tracked surgical pointer and the closest surface, while Lawonn et al. (2017) draw anchor lines between 3D rendered blood vessels and the internal surface of an interactively manipulated cylindrical cut-out region to allow users to get insight about the relative depth of different vessels.

A trivial approach to interaction that improves perception is to turn the visibility of virtual objects on and off. Choi et al. (2015) demonstrate that providing the possibility for a surgeon to interactively switch the view between AR and VR can improve the ability to reach a surgical target for spine surgery. These authors suggest that this is explained by a better depth perception in VR mode while the AR mode remains useful to provide context and be aware of the surface of the operating field. Bork et al. (2015) encode the distance of virtual objects to a tracked pointer in an AR scene by modulating their opacity following the temporal propagation of a spherical region of interest around the tip of the pointer. The progression of the region is also reflected in

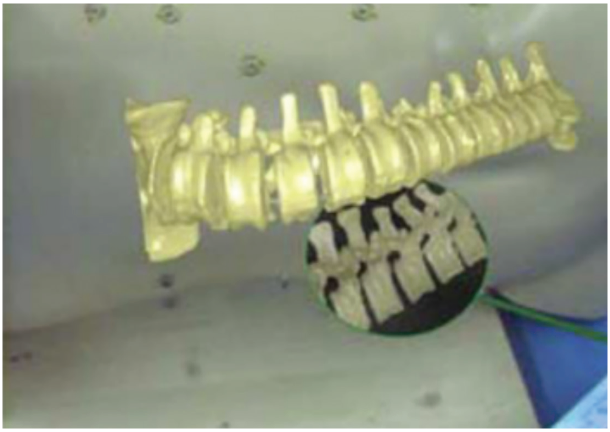

FIGURE 7.5 Example of the virtual mirror. (From Bichlmeier, C. et al., *IEEE Trans. Med. Imaging*, 28, 1498–510, 2009; Courtesy of Christoph Bichlmeier.)

an auditory display where a regular sound marks predefined steps of the progression, and a distinct sound marks the intersection of the region with virtual objects.

Controlling the point of view used to generate an AR image can greatly improve the perception. For all HMD-based systems, the feature is built-in, but the point of view is restricted by the surgeon's ability to move around the patient, which may be limited in a cluttered operating room or constrained by the use of a microscope. Shamir et al. (2006) propose to generate an AR view using a handheld tracked camera that enables fast exploration of the anatomy from various angles. Similarly, Kockro et al. (2009) employ a camera-equipped tracked surgical pointer in the OR to produce the augmented reality image. An alternative to interacting with the point of view is to use the virtual mirror, a concept inspired by a dentist's mirror, proposed by Bichlmeier et al. (2009). The interaction of the user with a tracked tool controls the 3D pose of a virtual mirror rendered in the AR scene and provides the ability to see behind 3D objects (c.f. Figure 7.5). Wieczorek et al. (2011) registered a preoperative 3D CT with an intraoperative x-ray, wherein depth information in the CT enabled the extrapolation of an x-ray image for a restricted range of depths. Various depth enhancement methods are proposed based on this principle. For example, tracked tools can be used to interactively manipulate a virtual plane used to "erase" parts of the x-ray beyond the plane.

7.4 DISCUSSION AND FUTURE DIRECTIONS

The development of novel interaction methods has been underrepresented in the AR IGS research community, with the majority of research focusing on the development of hardware, accurate and robust AR calibration techniques, and AR visualization methods. Yet, well-designed interaction techniques have the potential to simplify tasks, allow for more intuitive methods to control the IGS system and improve the surgical workflow, and enhance the perception of the guidance images and AR visualizations.

In terms of looking at interaction for *control*, some research groups have explored the use of touchless interaction, which has several advantages. These include the fact that the surgeon does not need to rely on someone else to interact with the system, which results in fewer misunderstandings between the surgeon and the person interacting with the IGS system. Further, this should lead to improved surgical workflows. Although there has been some research in this area, more work is needed to allow the surgeon to have easy, intuitive, and direct control over the IGS system.

Although not fully explored in the IGS literature, tangible user interfaces are another approach that may represent a novel solution to interactions for control in the surgical domain. Tangible user interfaces, which allow for the use of physical objects as a direct input mechanism for interaction with graphical representations, first became prominent with the work of Ishii and Ullmer (1997), and there have only been a handful of papers that have explored this type of interaction in medical imaging. Eck et al. (2016) developed a system for the preoperative planning phases of volume exploration and trajectory planning that uses a tracked handle to manipulate a preoperative volume and a force feedback stylus to re-slice the volume along an arbitrary plane. Hinckley et al. (1997) employed a tracked prop (plane or stylus), and a tracked head phantom, to explore an MRI by slicing the volume along the tracked prop axis. This field has yet to be fully explored in the IGS domain; however, one can envision what could be done in an OR, where more and more objects and devices will be modeled and tracked in real-time, and therefore may be used for interaction.

While there has been some work in terms of *task simplification*, we believe there is much potential for further research, particularly in terms of allowing the surgeon to use their knowledge to interact within the surgical field of view to account for inaccuracies of the system (such as brain shift, registration error, etc.), such as the work by Drouin et al. (2015). This would allow IGS systems to maintain accuracy and be used longer throughout surgery.

One of the main problems with AR in IGS is that the visualizations continue to be limited in terms of spatial and depth perception. Although this may be a minor issue in other domains, where labels and virtual elements may float above the real world, for IGS this is of utmost importance. Better depth perception of virtual elements between virtual anatomy and the surgical field of view and spatial relationships between these elements is needed for accurate guidance. By coupling interaction with visualization, it may be possible to enhance perception to allow for accurate localization of anatomy, and this will lead to a greater presence of AR in clinical practice.

REFERENCES

Abeele, S., and O. Bock. 2001. Mechanisms for Sensorimotor Adaptation to Rotated Visual Input. *Experimental Brain Research* 139, no. 2: 248–253.

Altenhoff, B.M., P.E. Napieralski, L.O. Long, J.W. Bertrand, C.C. Pagano, S.V. Babu, and T.A. Davis. 2012. Effects of Calibration to Visual and Haptic Feedback on near-Field Depth Perception in an Immersive Virtual Environment. In *Proceedings of the ACM Symposium on Applied Perception - SAP'12*, Vol. 71. New York: ACM Press.

Bajura, M., H. Fuchs, and R. Ohbuchi. 1992. Merging Virtual Objects with the Real World: Seeing Ultrasound Imagery within the Patient. *Computer Graphics* 26, no. 2: 203–210.

Bichlmeier, C., S.M. Heining, M. Feuerstein, and N. Navab. 2009. The Virtual Mirror: A New Interaction Paradigm for Augmented Reality Environments. *IEEE Transactions on Medical Imaging* 28, no. 9: 1498–1510.

Bichlmeier, C., T. Sielhorst, S.M. Heining, and N. Navab. 2007a. Improving Depth Perception in Medical AR A Virtual Vision Panel to the Inside of the Patient. In *Bildverarbeitung Für Die Medizin*, pp. 217–221. Berlin, Heidelberg: Springer.

Bichlmeier, C., F. Wimmer, S.M. Heining, and N. Navab. 2007b. Contextual Anatomic Mimesis Hybrid In-Situ Visualization Method for Improving Multi-Sensory Depth Perception in Medical Augmented Reality. In *IEEE/ACM International Symposium on Mixed and Augmented Reality*, pp. 1–10. IEEE.

Bier, E.A., M.C. Stone, K. Pier, W. Buxton, and T.D. DeRose. 1993. Toolglass and Magic Lenses. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH'93*, pp. 73–80. New York: ACM Press.

Bingham, G.P., and C.C. Pagano. 1998. The Necessity of a Perception-Action Approach to Definite Distance Perception: Monocular Distance Perception to Guide Reaching. *Journal of Experimental Psychology: Human Perception and Performance* 24, no. 1: 145–168.

Bork, F., B. Fuerst, and C. Graumann. 2015. Auditory and Visio-Temporal Distance Coding for 3-Dimensional Perception in Medical Augmented Reality. In *ISMAR 2015*.

Buckthought, A., A. Yoonessi, and C.L. Baker. 2014. Dynamic Perspective Cues Enhance Depth from Motion Parallax. *Journal of Vision* 14, no. 10: 734.

Choi, H., B. Cho, K. Masamune, M. Hashizume, and J. Hong. 2015. An Effective Visualization Technique for Depth Perception in Augmented Reality-Based Surgical Navigation. *The International Journal of Medical Robotics + Computer Assisted Surgery: MRCAS* (May 5).

Cutolo, F., G. Badiali, and V. Ferrari. 2015. Human-PnP: Ergonomic AR Interaction Paradigm for Manual Placement of Rigid Bodies. In *AE-CAI 2015, LNCS*, ed. C.A. Linte, Z. Yaniv, and P. Fallavollita, Vol. 9365, pp. 50–60. Cham, Switzerland: Springer.

Dergachyova, O., D. Bouget, A. Huaulmé, X. Morandi, and P. Jannin. 2016. Automatic Data-Driven Real-Time Segmentation and Recognition of Surgical Workflow. *International Journal of Computer Assisted Radiology and Surgery* 11, no. 6: 1081–1089.

Diotte, B., P. Fallavollita, L. Wang, S. Weidert, E. Euler, P. Thaller, and N. Navab. 2015. Multi-Modal Intra-Operative Navigation During Distal Locking of Intramedullary Nails. *IEEE Transactions on Medical Imaging* 34, no. 2: 487–495.

Dixon, B.J., M.J. Daly, H. Chan, A.D. Vescan, I.J. Witterick, and J.C. Irish. 2013. Surgeons Blinded by Enhanced Navigation: The Effect of Augmented Reality on Attention. *Surgical Endoscopy* 27, no. 2: 454–461.

Drascic, D., and P. Milgram. 1996. Perceptual Issues in Augmented Reality. In *SPIE Stereoscopic Displays and Virtual Reality Systems III*, pp. 123–134.

Drouin, S., M. Kersten-Oertel, and D.L. Collins. 2015. Interaction-Based Registration Correction for Improved Augmented Reality Overlay in Neurosurgery. In *Augmented Environments for Computer-Assisted Interventions*, Lecture Notes in Computer Science, Vol. 9365, pp. 21–29.

Eck, U., P. Stefan, H. Laga, C. Sandor, P. Fallavollita, and N. Navab. 2016. Exploring Visuo-Haptic Augmented Reality User Interfaces for Stereo-Tactic Neurosurgery Planning. In *Proceedings of the 7th International Conference on Medical Imaging and Virtual Reality, MIAR 2016*, pp. 208–220.

Fischer, J., J. Fischer, D. Bartz, and W. Straßer. 2005. Intuitive and Lightweight User Interaction for Medical Augmented Reality. In *Vision, Modeling and Visualization*, pp. 375–382.

Forestier, G., F. Lalys, L. Riffaud, D. Louis Collins, J. Meixensberger, S.N. Wassef, T. Neumuth, B. Goulet, and P. Jannin. 2013. Multi-Site Study of Surgical Practice in Neurosurgery Based on Surgical Process Models. *Journal of Biomedical Informatics* 46, no. 5: 822–829.

Gras, G., H.J. Marcus, C.J. Payne, and P. Pratt. 2015. Visual Force Feedback for Hand-Held Microsurgical Instruments. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Vol. 9349, pp. 480–487.

Gratzel, C., T. Fong, S. Grange, and C. Baur. 2004. A Non-Contact Mouse for Surgeon-Computer Interaction. *Technology & Health Care* 12, no. 3: 245–257.

Gregory, R.L. 1977. *Eye and Brain, the Psychology of Seeing*, 3rd ed. London, UK: Weidenfeld & Nicolson.

Harman, K.L., G.K. Humphrey, and M.A. Goodale. 1999. Active Manual Control of Object Views Facilitates Visual Recognition. *Current Biology* 9, no. 22: 1315–1318.

Harris, C.S. 1965. Perceptual Adaptation to Inverted, Reversed, and Displaced Vision. *Psychological Review* 72, no. 6: 419–444.

Herrlich, M., P. Tavakol, D. Black, D. Wenig, C. Rieder, R. Malaka, and R. Kikinis. 2017. Instrument-Mounted Displays for Reducing Cognitive Load During Surgical Navigation. *International Journal of Computer Assisted Radiology and Surgery* 12, no. 9: 1599–1605.

Hinckley, K., R. Pausch, J. Hunter Downs, D. Proffitt, and N.F. Kassell. 1997. The Props-Based Interface for Neurosurgical Visualization. *Studies in Health Technology and Informatics*, vol. 39: 552–562.

Ishii, H., and B. Ullmer. 1997. Tangible Bits. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '97*, pp. 234–241.

Jannin, P., M. Raimbault, X. Morandi, L. Riffaud, and B. Gibaud. 2003. Model of Surgical Procedures for Multimodal Image-Guided Neurosurgery. *Computer Aided Surgery* 8, no. 2: 98–106.

Johnston, E.B., B.G. Cumming, and M.S. Landy. 1994. Integration of Stereopsis and Motion Shape Cues. *Vision Research* 34, no. 17: 2259–2275.

Kalkofen, D., E. Mendez, and D. Schmalstieg. 2009. Comprehensible Visualization for Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics* 15, no. 2: 193–204.

Kantelhardt, S.R., A. Gutenberg, A. Neulen, N. Keric, M. Renovanz, and A. Giese. 2015. Video-Assisted Navigation for Adjustment of Image-Guidance Accuracy to Slight Brain Shift. *Neurosurgery* 11, no. 4: 1–8.

Katić, D., G. Sudra, S. Speidel, G. Castrillon-Oberndorfer, G. Eggers, and R. Dillmann. 2010. Knowledge-Based Situation Interpretation for Context-Aware Augmented Reality in Dental Implant Surgery. In *Medical Imaging and Augmented Reality, MIAR 2010*, pp. 531–540. Berlin, Germany: Springer.

Kersten-Oertel, M., I. Gerard, S. Drouin, K. Mok, D. Sirhan, D.S. Sinclair, and D.L. Collins. 2015. Augmented Reality in Neurovascular Surgery: Feasibility and First Uses in the Operating Room. *International Journal of Computer Assisted Radiology and Surgery* 10, no. 11: 1823–1836.

Kersten-Oertel, M., P. Jannin, and D.L. Collins. 2013. The State of the Art of Visualization in Mixed Reality Image Guided Surgery. *Computerized Medical Imaging and Graphics* 37, no. 2: 98–112.

Kirmizibayrak, C., N. Radeva, M. Wakid, J. Philbeck, J. Sibert, and J. Hahn. 2011. Evaluation of Gesture Based Interfaces for Medical Volume Visualization Tasks. In *Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry - VRCAI'11*, Vol. 69. New York: ACM Press.

Klatzky, R.L., J.M. Loomis, A.C. Beall, S.S. Chance, and R.G. Golledge. 1998. Spatial Updating of Self-Position and Orientation During Real, Imagined, and Virtual Locomotion. *Psychological Science* 9, no. 4: 293–298.

Knill, D.C. 2005. Reaching for Visual Cues to Depth: The Brain Combines Depth Cues Differently for Motor Control and Perception. *Journal of Vision* 5, no. 2: 2.

Kocev, B., F. Ritter, and L. Linsen. 2014. Projector-Based Surgeon–computer Interaction on Deformable Surfaces. *International Journal of Computer Assisted Radiology and Surgery* 9, no. 2: 301–312.

Kockro, R.A., Y.T. Tsai, I. Ng, P. Hwang, C. Zhu, K. Agusanto, L.X. Hong, and L. Serra. 2009. Dex-Ray: Augmented Reality Neurosurgical Navigation with a Handheld Video Probe. *Neurosurgery* 65, no. 4: 795–808.

Kutter, O., A. Aichert, C. Bichlmeier, J. Traub, S.M. Heining, B. Ockert, E. Euler, N. Navab, and S. Michael. 2008. Real-Time Volume Rendering for High Quality Visualization in Augmented Reality. In *International Workshop on Augmented Environments for Medical Imaging Including Augmented Reality in Computer-Aided Surgery.*

Lalys, F., and P. Jannin. 2014. Surgical Process Modelling: A Review. *International Journal of Computer Assisted Radiology and Surgery* 9, no. 3: 495–511.

Landy, M.S., L.T. Maloney, E.B. Johnston, and M. Young. 1995. Measurement and Modeling of Depth Cue Combination: In Defense of Weak Fusion. *Vision Research* 35, no. 3: 389–412.

Lawonn, K., M. Luz, and C. Hansen. 2017. Improving Spatial Perception of Vascular Models Using Supporting Anchors and Illustrative Visualization. *Computers & Graphics* 63: 37–49.

Lerotic, M., A.J. Chung, G. Mylonas, and G.-Z. Yang. 2007. Pq-Space Based Non-Photorealistic Rendering for Augmented Reality. *International Conference on Medical Image Computing and Computer-Assisted Intervention* 10, no. Pt 2: 102–109.

Marotta, J.J., T.S. Perrot, D. Nicolle, P. Servos, and M.A. Goodale. 1995. Adapting to Monocular Vision: Grasping with One Eye. *Experimental Brain Research* 104, no. 1: 107–114.

Masia, L., M. Casadio, G. Sandini, and P. Morasso. 2009. Eye-Hand Coordination during Dynamic Visuomotor Rotations. *PloS One* 4, no. 9: e7004.

Mendez, E., D. Kalkofen, and D. Schmalstieg. 2006. Interactive Context-Driven Visualization Tools for Augmented Reality. In *IEEE/ACM International Symposium on Mixed and Augmented Reality*, pp. 209–218. IEEE.

Mewes, A., B. Hensen, F. Wacker, and C. Hansen. 2017. Touchless Interaction with Software in Interventional Radiology and Surgery: A Systematic Literature Review. *International Journal of Computer Assisted Radiology and Surgery* 12, no. 2: 291–305.

Navab, N., J. Traub, T. Sielhorst, M. Feuerstein, and C. Bichlmeier. 2007. Action- and Workflow-Driven Augmented Reality for Computer-Aided Medical Procedures. *IEEE Computer Graphics and Applications* 27, no. 5: 10–14.

Onceanu, D., and A.J. Stewart. 2011. Direct Surgeon Control of the Computer in the Operating Room. In *Medical Image Computing and Computer-Assisted Intervention*, pp. 121–128. Berlin, Germany: Springer.

Pauly, O., B. Diotte, P. Fallavollita, S. Weidert, E. Euler, and N. Navab. 2015. Machine Learning-Based Augmented Reality for Improved Surgical Scene Understanding. *Computerized Medical Imaging and Graphics* 41: 55–60.

Rieser, J.J., D.A. Guth, and E.W. Hill. 1986. Sensitivity to Perspective Structure While Wa/King without Vision. *Perception* 15, no. 1: 173–188.

Rong Wen, C.-B. Chng, and C.-K. Chui. 2017. Augmented Reality Guidance with Multimodality Imaging Data and Depth-Perceived Interaction for Robot-Assisted Surgery. *Robotics* 6, no. 2: 13.

Salb, T., J. Brief, T. Welzel, B. Giesler, S. Hassfeld, J. Muehling, and R. Dillmann. 2003. INPRES (Intraoperative Presentation of Surgical Planning and Simulation Results): Augmented Reality for Craniofacial Surgery. In *SPIE Electronic Imaging: Stereoscopic Displays and Virtual Reality Systems*, ed. A.J. Woods, M.T. Bolas, J.O. Merritt, and S.A. Benton, Vol. 453. International Society for Optics and Photonics.

Seitel, A., N. Bellemann, M. Hafezi, A.M. Franz, M. Servatius, A. Saffari, T. Kilgus et al. 2016. Towards Markerless Navigation for Percutaneous Needle Insertions. *International Journal of Computer Assisted Radiology and Surgery* 11, no. 1: 107–117.

Shamir, R., L. Joskowicz, and Y. Shoshan. 2006. An Augmented Reality Guidance Probe and Method for Image-Guided Surgical Navigation. In *5th International Symposium on Robotics and Automation*, pp. 1–6.

Shamir, R.R., M. Horn, T. Blum, J.H. Mehrkens, Y. Shoshan, L. Joskowicz, and N. Navab. 2011. Trajectory Planning with Augmented Reality for Improved Risk Assessment in Image-Guided Keyhole Neurosurgery. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 1873–1876.

Shimamura, S., M. Kanegae, Y. Uema, M. Inami, T. Hayashida, H. Saito, and M. Sugimoto. 2013. Virtual Slicer: Development of Interactive Visualizer for Tomographic Medical Images Based on Position and Orientation of Handheld Device. In *2013 International Conference on Cyberworlds*, pp. 383–383. IEEE.

Simons, D.J., R.F. Wang, and D. Roddenberry. 2002. Object Recognition Is Mediated by Extraretinal Information. *Perception & Psychophysics* 64, no. 4: 521–530.

State, A., M.A. Livingston, W.F. Garrett, G. Hirota, M.C. Whitton, E.D. Pisano, and H. Fuchs. 1996. Technologies for Augmented Reality Systems. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH'96*, pp. 439–446. New York: ACM Press.

Sudra, G., S. Speidel, D. Fritz, B.P. Müller-Stich, C. Gutt, and R. Dillmann. 2007. MEDIASSIST: Medical Assistance for Intraoperative Skill Transfer in Minimally Invasive Surgery Using Augmented Reality. In *SPIE Medical Imaging 2007*, ed. K.R. Cleary and M.I. Miga. International Society for Optics and Photonics.

Wen, R., W.L. Tay, B.P. Nguyen, C.B. Chng, and C.K. Chui. 2014. Hand Gesture Guided Robot-Assisted Surgery Based on a Direct Augmented Reality Interface. *Computer Methods and Programs in Biomedicine* 116, no. 2: 68–80.

Wexler, M., and J.J.A. van Boxtel. 2005. Depth Perception by the Active Observer. *Trends in Cognitive Sciences* 9, no. 9: 431–438.

Wieczorek, M., A. Aichert, P. Fallavollita, O. Kutter, A. Ahmadi, L. Wang, and N. Navab. 2011. Interactive 3D Visualization of a Single-View X-Ray Image. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2011 SE - 10*, Lecture Notes in Computer Science. ed. G. Fichtinger, A. Martel, and T. Peters, Vol. 6891, pp. 73–80. Berlin, Germany: Springer.

Wipfli, R., V. Dubois-Ferrière, S. Budry, P. Hoffmeyer, C. Lovis, J. Wobbrock, and W. Schlegel. 2016. Gesture-Controlled Image Management for Operating Room: A Randomized Crossover Study to Compare Interaction Using Gestures, Mouse, and Third Person Relaying. Ed. Peter M.A. van Ooijen. *PLoS One* 11, no. 4: e0153596.