Abstract
This paper presents a review of the state of the art of visualization in mixed reality image guided surgery (IGS). We used the DVV (Data, Visualization processing, View) taxonomy to classify a large unbiased selection of publications in the field. The goal of this work was not only to give an overview of current visualization methods and techniques in IGS but more importantly to analyze the current trends and solutions used in the domain. In surveying the current landscape of mixed reality IGS systems, we identified a strong need to assess which of the many possible data sets should be visualized at particular surgical steps, to focus on novel visualization processing techniques and interface solutions, and to evaluate new systems.

1. Introduction
In image-guided surgery (IGS), surgical instruments are tracked and visualized with respect to patient-specific data sets to guide the surgeon in their tasks. IGS has driven the field of minimally invasive surgery (MIS), making exploratory surgery all but extinct and open invasive surgery less common. The advantages of MIS include smaller incisions, faster recovery times and better patient outcomes with lower morbidity and mortality rates. With this move to less invasive surgery, however, came the need for novel visualization techniques to overcome the lack of a direct view of the patient anatomy and the surgeons limited field of view. Mixed reality visualizations have offered one solution to this problem.

Mixed reality is the area on the reality-virtuality continuum [1] between reality, the unmodeled real environment, and virtual reality (VR), a purely virtual and modeled environment. Where an environment lies on the continuum corresponds to the extent to which the environment is modeled and whether real or virtual objects are introduced into it. Both augmented reality (AR) and augmented virtuality (AV), two points on the reality-virtuality continuum, have been used in IGS to visualize preoperative and intraoperative models of the patient registered with the real environment. These types of visualizations allow the surgeon a more extensive view beyond the visible anatomical surface of the patient and a better understanding of the spatial relationship between the anatomical models and real anatomy. Such an understanding could reduce surgical time, increase surgical precision and enable a better intraoperative understanding of the patient anatomy.

In this paper, we present a literature review based on the DVV (Data, Visualization processing, View) taxonomy [2] of selected papers in the domain of IGS. We aim to give an exhaustive overview of current mixed reality visualization techniques used clinically in IGS. The DVV taxonomy enables a systematic evaluation of each component of a mixed reality IGS system. Such a systematic analysis of the major entities of current clinical systems will allow future developers to more easily recognize which components can be reused and which need improvement or new solutions. Furthermore it may aid developers and end-users discuss and understand the constituents of a mixed reality visualization system, facilitating a greater presence of future systems in the OR.

The remainder of the paper is organized as follows: Section 2 describes the criteria used to create the database of reviewed publications and, briefly, the DVV taxonomy used to classify the selected publications. Section 3 provides a review of the papers organized by each of the components of the DVV taxonomy. An analysis of the papers which aims at giving a better understanding of the current solutions available in the area of mixed reality visualization systems for IGS and at identifying any commonalities across systems both within and across surgical domains is given in Section 4. Conclusions are given in Section 5.

2. Methodology
This section describes our methodology for reviewing mixed reality IGS systems, beginning with the search method used to create the database of reviewed papers and followed by a description of the DVV taxonomy [2] used to classify the selected publications.

2.1. Search Method
To create an unbiased selection of relevant state-of-the-art publications that describe mixed reality visualizations systems for IGS, we conducted a query for publications with the terms reality or virtuality in the title along with terms related to surgery and excluding terms related to planning, education or training. Although this query does not produce an exhaustive search of papers for review, it does allow for a systematic and subjective method of selecting papers. The limitations of the query method are further discussed in Section 4.
The following two advanced Google Scholar searches (last updated on August 25, 2011), limited to search within Biology, Life Sciences, and Environmental Science Engineering, Computer Science, and Mathematics Medicine, Pharmacology, and Veterinary Science journals, were done:

1. allintitle: reality CAS OR "computer assisted" OR medical OR surgery OR surgical OR intraoperative OR "intra-operative" OR neurosurgery OR operating -planning -education -simulation -simulations -simulator -simulators -trainer -training -myth -plastic

2. allintitle: virtuality CAS OR "computer assisted" OR medical OR surgery OR surgical OR intraoperative OR "intra-operative" OR neurosurgery OR operating -planning -education -simulation -simulations -simulator -simulators -trainer -training -myth -plastic

The first search returned 467 articles, and the second, 3 articles. Of these, 87 articles met our inclusion criteria as follows. The publications had to include a description of an entire visualization system for clinical use in the operating room (OR), although the focus could be on only one aspect of the system, for example, calibration, image processing, display technology and so on. Furthermore, the publications had to describe a system for use in the OR rather than a simulator for diagnosis, planning or training.

2.2. The DVV Taxonomy: Classification for Mixed Reality Visualization in IGS

We used the DVV taxonomy to organize our review of the state of the art in mixed reality visualization in IGS. The DVV taxonomy describes the three main factors that should be considered when developing a system in order to provide surgeons with an effective tool to perform surgery: 1) which of the abundance of available data should be used and shown to the end user; 2) how the data can be effectively merged and visualized; and 3) how it can be best displayed, viewed and interacted with. A full description of the DVV taxonomy can be found in Kersten-Oertel et al. [2].

By using the terms and classes that constitute the DVV taxonomy to examine the state of the art in mixed reality IGS systems, we aim to provide developers with an overview of the current systems in the domain and a better understanding of how to implement new systems and components for clinical use.

3. State of the art using the DVV taxonomy

This section describes how each of the DVV components is addressed in the 87 publications in our database. To allow the reader to refer to the definitions of the DVV taxonomy, the section numbering used here is consistent with the numbering used in the publication describing the DVV taxonomy [2]. For example, in the DVV taxonomy paper, section 3.1.1 explains patient-specific data, and here, section 3.1.1 describes what type of patient-specific data was used in the selected publications. A table showing the classification of the selected publications according to the DVV taxonomy is available online at http://www.bic.mni.mcgill.ca/Research/Labs/PLDVV/. A deeper analysis and interpretation of the consequences of choosing a particular solution for each of the components is provided in Section 4.

3.1. Data

The first component of the DVV taxonomy, data, has two main superclasses: patient-specific data and visually processed data. The use of data as a factor highlights the importance of selecting which of the many available data sets for a given patient should be shown to the end user and at which point during the surgery.

3.1.1. Patient-Specific Data

Although patient-specific data includes demographics and clinical scores, the focus here is on signal or raw imaging data and analyzed imaging data. Raw imaging data is the direct output of the acquisition system. It is transformed into analyzed imaging data and only then visualized. A visualization depends on a few main attributes of the raw and analyzed data: the dimensionality of the acquisition sensor, whether the data is preoperative or intraoperative and the data primitive of the analyzed imaging data used.

In terms of raw imaging data, in most papers, the system (and the surgical task) relies on the use of either preoperative computed tomography (CT) or magnetic resonance imaging (MRI) data sets of the patient in combination with real-time video from either a camera or an endoscope. A few publications identified intraoperative data from X-ray fluoroscopy [3], ultrasound [4, 5, 6], MRI [6], CT [7] or transesophageal echocardiography (TEE) [8, 9]. Nine of the publications [10, 11, 12, 13, 14, 15, 16, 17, 18] did not specify the raw imaging data.

3.1.2. Visually Processed Data

Visually processed data is the output of visualization processing; it is presented to the end user by means of the view. Three subclasses of visually processed data are defined in the DVV taxonomy: analyzed imaging data, derived data and prior knowledge data.

The main attribute considered for visually processed data is its semantic, defined as its meaning at a particular surgical step. The semantic may be anatomical (dealing with the anatomy, physiology or pathology of the patient), operational (dealing with actions or tasks) or strategic (dealing with planning and guidance). The majority of systems described in the papers use an anatomical semantic; that is, models of the patients anatomy are visualized to aid the surgeon in localizing and navigating to a target.

Six of the systems [19, 20, 13, 21, 22, 23, 24] visualize data with an operational semantic. This data is generally in the form of the distance from a surgical instrument to a target and is shown either numerically, visually (using color or state transition) or in both ways. Kawamata et al.’s [21] navigation system
for transsphenoidal surgery shows the distance from the surgical tool to the tumor in a bar graph. Similarly, Soler et al.’s AR laparoscopic system [22, 23] displays the numerical distance to a tumor, and the surgical needle changes color when pointing at the target [23]. In Trevisan et al.’s [24] AR system for maxillofacial surgery, a dynamic sphere shows the distance of the tool from the dental nerve. Furthermore, the virtual representation of the tool turns green when on the correct planned path. The system by Birkfellner et al. [19] visually conveys distance between the surgical tool and the target by changing the representation of the target from solid (or surface rendered) to wireframe. Katić et al. [13] described the visualization of distances between objects using color-coding based on distance information. Giraldez et al. [20] gave the user the option to compute and display the distance between the surgical tools and a target anatomy.

Several papers described systems in which data with a strategic semantic is visualized. In Wagner et al.’s [25] orthognathic AR navigation system virtual objects include yellow lines representing the planned osteotomy and a blue line representing the saw tool. The neurosurgical navigation system proposed by Wörn et al. [26] visualizes planned trajectories for bone cutting, boreholes and biopsy entry points. Seven publications specified visualization of the extrapolated path or trajectory of surgical tools [20, 27, 28, 14, 29, 30, 18]. Although it is likely that planning data is also visualized in other systems, no other publications specified the visualization of such data.

In the following sections, the subclasses of visually processed data (analyzed imaging, prior knowledge and derived data) are addressed individually.

**Analyzed Imaging Data** The main attribute of analyzed imaging data is the underlying data primitive used by the system. Instances include point, line, contour, plane, surface, wireframe and volume. In 21 of the 87 selected publications, the authors did not indicate the type of data primitive used. The remainder of the papers specified one or more data primitives to represent the virtual objects in the visualization.

Six papers [25, 31, 32, 33, 34, 35] described systems that use points in combination with some other data primitives. In the projector system developed by Glossop and colleagues [31, 32, 33], only points and lines or contours can be used to delineate the boundary of a craniotomy. Wagner et al.’s system for orthognathic surgery [25] uses colored points to mark the target (lesion or tumor). Suzuki et al. [35] use a large on-screen point indicator to convey the hardness of an object grasped by the robotic manipulator (green for soft material; yellow and red for hard).

Typically, lines and contours are used to depict margins of objects of interest [36, 37], plans [25, 38, 39] or tool paths [38, 39, 25]. Some developers employ texture-mapped planes to show either two-dimensional (2D) data, such as X-ray fluoroscopic images [3], X-ray images [40] or 2D ultrasound [41, 42, 20], or slices of three-dimensional (3D) CT or MRI data [43, 44, 30]. Wireframe or mesh representations of a patients anatomy are specified in 11 of the selected publications [21, 45, 37, 46, 29, 28, 38, 19, 47, 48, 49].

The systems in almost half of the publications use surfaces by themselves or in combination with another data primitive. In these systems, surfaces are typically texture-mapped color-coded anatomical models obtained by segmenting raw imaging data. Some systems, however, also use surfaces to represent virtual surgical instruments [50, 20, 27, 8, 43, 9, 22, 51, 52, 30, 53], targets or markers [19, 37, 10], or prosthetics and implants [43, 9, 42, 40].

Ten publications described the use of volumetric objects to represent the analyzed data [54, 5, 55, 56, 57, 58, 35, 59, 53, 60]. In Spletchna et al.’s [6] system volume-rendered intraoperative ultrasound scans are visualized along with surface-rendered vessel trees for liver surgery.

**Prior Knowledge Data** Prior knowledge data refers to data derived from generic models, for example, atlases, prior measurements or plans, tool models, uncertainty information specific to the IGS system and surgical road maps. Less than one-third of the publications described visualization of prior knowledge data, and most of these did so only in terms of tool depiction. The tools localized and displayed on screen to guide the surgeon include an endoscope [21], saw [25], surgical needle [47, 30], surgical drill [53], Raman probe [34], ultrasound probe [61], surgical instrument arms for cardiovascular surgery [51], endovascular tools [42] and general surgical instruments [50, 10, 11, 12, 20, 62, 49, 63, 27, 8, 9, 64, 15, 65]. A number of systems also visualize markers, prosthetics and implants [61, 37, 43].

Prior knowledge data also includes planning data that may be transferred to the anatomical data to guide the surgeon. Only three publications explicitly specified planning data as visually processed data. The neuronavigation system of Wörn et al. [26] visualizes boreholes, trajectories for bone cutting and biopsy points. Salb et al.’s [14] head mounted display (HMD) craniofacial navigation system allows the user to choose to render the preoperative planning data of the intervention, and Wagner et al.’s [25] system displays the planned osteotomy lines to the end user. The rest of the selected publications made no mention of whether prior knowledge data is used and, if so, how it is rendered on screen.

**Derived Data** Derived data is obtained from processing either patient-specific data only or both patient-specific data and prior knowledge data. Examples of derived data include labels, uncertainty information specific to the patient or type of surgery and intraoperative measurements such as distances between regions of interest (ROIs).

Six of the publications described the visualization of derived data in terms of the distance from a tool to the target anatomy. Kawamata et al.’s [21] system shows the distance numerically in a bar graph; Soler et al. [22, 23] and Giraldez et al. [20] show it numerically; and Trevisan et al. [24], with a color-changing dynamic sphere. The system by Birkfellner et al. [19] changes the representation of the target from solid to wireframe based on the distance measurement. In Katić et al. [13], distances between objects are depicted with a color-coding scheme.

A number of systems also visualize derived data in terms of the computed trajectory or extrapolated tool path. Extrapolated needle paths are shown in the systems by Sauer et al. [29], Wen et al. [18] and Vogt et al. [30]. Tomikawa et al.’s [47] sys-
tem depicts the puncture line of the surgical needle in breast-conserving surgery, and two publications [28, 20] describe the visualization of other tool trajectories.

Other types of derived data are also visualized in different systems. Sato et al.’s [66] breast cancer surgical navigation system depicts measurements of the cancers intraductal spread; red marks indicate intraductal spread and green if no intraductal spread is found. Linte et al.’s [43] system graphically represents uncertainty information in terms of target registration error (TRE) using a 95% confidence ellipsoid. In Salb et al.’s [14] AR craniotomy system, risk distributions are calculated with an isotropic risk potential and anisotropic tissue field to determine the access paths with the least risk. Suzuki et al.’s [35] surgical robot with AR functions displays haptic sense information from the right and left robot arms, the forceps indicator, the location of the robot tip, and the patients vital signs. In the liver guidance system described by Splechtna et al., [6], the radiologist intraoperatively places annotation markers (in the form of colored spheres) on the vessel tree to note ROIs for the surgeon. King et al. [34] represent different tissue classes with colored markers. Zheng et al.’s [40] long bone fracture navigation system computes and visualizes the outlines and the centerlines of bone fragments. Volonté et al. [59] use arrows pointing from the portals of the laparoscopic tools to the target to show a visual triangulation with the laparoscopic instruments in the system. In their AR system for arthroscopic knee surgery, Tonet et al. [46] place a red highlighted area on the anatomical surface model to represent the area seen in the arthroscopic view. Wen et al.’s [18] needle guidance system for RF tumor ablation visualizes the region of necrosis resulting from the insertion and location of the needle.

In the publications that described the visualization of derived data, the data is typically in the form of patient-specific measurements [66, 35, 6, 18, 40] or uncertainty information specific to the surgery or the patient [43, 14]. Such additional information is intended to aid the surgeon in making decisions throughout the surgical process.

3.2. Visualization Processing

The second component of the DVV taxonomy is visualization processing. Visualization processing describes the specific techniques or transformations performed on the data to achieve the best visual representation of the data for a particular task at a given surgical step. Surprisingly, over one-third of the publications did not mention what type of visualization processing is performed.

In the selected publications, visualization processing deals mostly with the visualization of anatomical data and is limited to color-coding structures [67, 54, 13, 21, 34, 68, 43, 69, 70, 71, 72, 9, 66, 22, 23, 29, 57, 44, 65, 73, 35, 47, 74, 24, 59, 25, 53], using saliency or emphasis techniques to highlight ROIs [50, 54, 57, 71, 22, 46, 59, 53, 75] and using transparency to combine modality data [67, 54, 76, 68, 77, 27, 43, 69, 3, 72, 9, 48, 57, 15, 7, 22, 23, 73, 35, 24, 74, 30, 75]. Most of the systems use simple rendering techniques, such as color-coding and transparency.

A common problem with mixed reality visualization is the perception that the virtual object lies above the real object rather then below its surface. Bichlmeier et al. [50] proposed a more sophisticated way of using transparency that improves depth perception. In their work, the transparency function is based on the skin curvature as well as the observer’s line of sight, resulting in a more sophisticated visualization of the virtual object. The end effect is a better perception of the depth at which the virtual objects lay within the real scene.

It is often difficult to perceive the 3D shape and structure of rendered medical data. Although the cues humans use to perceive depth in everyday life have been thoroughly studied, what cues and how they should be combined to convey depth in computer-generated images remains an ongoing topic of research. In mixed reality visualization, not only are shape and structure important, but so is the relative depth between objects. Only a few groups have gone beyond using transparency alone and studied the use of different perceptual cues to enhance the proper perception of depth of virtual objects in a real scene or real objects in a virtual scene.

Fuchs et al. [78] use the depth cue of occlusion to aid in the perception of the order of objects. Specifically, a pixel is painted only if it lies above the surface of the closest object. Samset et al. [48] use edges extracted from the real video stream to provide a depth cue, helping the user perceive the overlaid virtual object as being within, rather than floating in front of, the patient. In Wieczorek et al.’s [53] AR system, occlusion handling of the surgeon’s hand gives a more realistic percept. In Wimmer et al. [75], an extra linear gradient texture (where light was closer to the viewer and dark further away) improves depth perception. The AR system developed by Edwards et al. [36] incorporates hidden line removal to help with depth perception and also enables stereo projection to allow perception of the virtual object at the correct distance below the patient surface. Suzuki et al. [73] employ the cue of stereopsis in the form of stereo pairs. The superimposed anatomy (virtual object) is shown in two windows; the top view shows the left-eye image and the bottom view, the right-eye image. Bichlmeier et al. [50] system allows the user to make use of the motion parallax cue by moving their head and looking at the patient from different points of view. Liao et al. [63] developed an AR system for MRI-guided procedures using integral videography (IV) and a half-silvered mirror. In their IV system, a 3D object is reconstructed from light sources such that each 3D point in space represents the convergence of rays from a number of pixels from the computer monitor and therefore appears as a new light source [63]. Because geometrically accurate stereoscopic images can be viewed from various directions, the surgeon may also exploit the motion parallax cue.

A number of surgical systems apply lighting and shading cues, which help determine an object’s shape as well as its spatial position, to render virtual objects. Four systems use shading such as Phong shading as a depth cue [50, 57, 37, 53]. Particularly notable is the visualization by Bichlmeier et al. in which the anatomical surface models and the surgical instruments cast realistic shadows on each other, providing depth feedback to aid navigation of the surgical instrument [50].

State visualization is another way of visualizing data that helps the user determine when the context or the environment
has changed. In surgical navigation systems, it is common to visualize state changes. by using color or rendering by changing the representation of the analyzed data object. As mentioned in Section 3.1.2, a number of systems use state visualization by changing the color [21, 30, 29] or rendering style [19] of the analyzed data object.

Lastly, one publication [40] described the use of non-photorealistic (NPR) rendering in a computer-assisted AR virtual fluoroscopic surgical system for long bone fractures. In the initial system, photorealistic rendering caused the implant models to occlude important aspects of the fluoroscopic images. Zheng et al. [40] improved the system by using a more illustrative style for surgical implants and found that such a visualization provides more image details and does not occlude the fluoroscopic image during navigation.

3.3. View

View, the third factor of the DVV taxonomy, comprises three main components: perception location, display and interaction tools. Perception location is the area of the environment on which the user must focus attention to view the mixed reality visualization. Display is the particular technology used to present visualizable data to the end user. Lastly, we have the interaction tools which deal with the design, implementation and evaluation of computer systems used by humans. We focus on two major subclasses of interaction tools compose the system’s user interface: hardware and virtual interaction tools. Hardware interaction tools are the physical devices the user employs to manipulate data, whereas virtual interaction tools allow the user to manipulate the pose and the view of the visualizable data as well as the datas visualization parameters.

3.3.1. Perception Location

Just under half of the papers (43) identified the perception location as the patient. Three papers [76, 77, 79] did not specify the perception location, and in three other systems [43, 72, 48], the perception location could be either the patient or the monitor. In the remainder of the systems, the perception location is a monitor.

In the endoscopic surgical systems [5, 73, 80], including laparoscopic [67, 70, 58, 7, 22, 74, 59] and arthroscopic systems [46], the perception location is almost always the monitor, with few exceptions. This is perhaps the case as an external monitor is traditionally used in these types of surgery. In Fuch’s et al.’s [78] laparoscopic system, the camera image is augmented with wireframe preoperative models and displayed to the end user via an HMD. In their thoracoscopic AR system, Sauer et al. [64] also use an HMD to depict surface models of the spine and objects of interest on the patient, with the thoracoscopic view inset in the end users AR view. Paolis et al. [70] presented a laparoscopic system for abdominal surgery wherein the preoperative surface models of vessels and ROIs are displayed on the patient, although the display hardware is not specified. Sudra et al.’s system [16] uses an HMD to augment reality in endoscopic robotic surgery.

With the exception of four papers [68, 69, 3, 71], in all publications that discussed neuronavigation systems, the patient is the perception location. In the work by Paul et al. [72], the perception location can be either the patient, using the microscope as the display device, or a monitor.

The trend of choosing one perception location over another for a particular type of surgery suggests that perception location may be closely linked with the surgical domain. This issue is more closely examined in Section 4.

3.3.2. Display

Display type was not specified in five papers, either because the work is preliminary [76], the system is generic and any instance of a component can be used [48], or the authors simply did not mention it [77, 79, 70]. In the rest of the publications, the display used varies, with the perception location being either the patient or an external display device.

The systems in over one-third of the publications use some form of monitor. Three papers specified a live video monitor (displaying a live video feed) [68, 69, 58]; three, an autostereoscopic monitor [81, 62], and the remainder, generic display monitors [21, 66, 3, 51, 22, 46, 27, 71, 5, 61, 23, 28, 42, 34, 57, 35, 73, 67, 9, 8, 80, 49, 7, 44, 40, 74, 13]. In Linte et al.’s [43] cardiac surgery system, either an HMD for AR or a monitor for AV can be used. Similarly, Lo et al.’s [41] cardiac navigation system allows either an HMD or a monitor as the display device. In the neuronavigation platform proposed by Paul et al. [72], the display device can be either a monitor or the microscope.

In publications where the patient is the perception location, a wide variety of display devices were specified: A few systems used a surgical microscope or a head-mounted variation of one. At Imperial College London, researchers developed an AR system for microscope-assisted guided interventions (MAGI) [36, 45, 82, 37] that uses a stereo microscope as the display device. The ability to project a different image into each eyepiece enables correct stereoscopic visualization of virtual features presented in the optical path of the microscope. In a similar effort, based on a commercially available head-mounted operating binocular, Birkfellner et al. [83, 84, 10, 85] developed the Varioscope AR: a custom-built, head-mounted operating microscope. The Varioscope AR, which is outfitted with two miniature VGA displays, allows the projection of virtual objects onto the focal plane of the main lens. Similarly, Aschke et al. [4] outfitted an operating microscope with a microscopic bench connected to the microscope’s beam splitters to allow for correct stereoscopic visualization of the virtual features presented in the optical path of the microscope.

Mischkowski et al. [55, 56] use a portable LCD screen, the X-Scope. The screen is equipped with a digital camera on the back, and an attached reference is tracked using infra-red cameras. Given the spatial orientation of the screen, preoperative virtual models are aligned and superimposed on the real image of the same object as captured by the digital camera. The advantage of this device lies in the surgeon’s ability to hold the screen and walk around and look at the patient from various points of view with completely natural interaction.

The ability to move freely around the OR and see the patient from multiple points of view is also an advantage of using
an HMD. The systems described in 17 papers use an HMD, which may be video see-through or optical see-through, to display the mixed reality scene to the end user. In video see-through HMDs, two video cameras are attached to the head gear, and the captured video images of the real world are combined with computer-generated objects to be viewed by the user. In optical see-through HMDs, optical mirrors are used to reflect computer-generated images above the real world. Another possibility is that a purely virtual environment is viewed using the HMD with only computer-generated images, however, this type of system was not described in the publications. Most of the systems that use HMDs, among the selected papers, employ video see-through systems [78, 50, 54, 86, 41, 29, 64, 16, 30, 53, 75]. Only one system uses an optical see-through [14] HMD. In three papers [25, 43, 6], the authors did not specify the type of HMD. Yamaguchi et al. [60] use a retinal projection HMD, which projects images directly onto the eye rather than on an external display. Suthau et al. [65] examined different AR visualization solutions in the context of liver surgery and compared video see-through HMDs, optical see-through HMDs and virtual retinal displays/retinal projection HMDs. The authors concluded that there is a need for better see-through technologies, and that current displays are suboptimal. Specifically, they cited HMDs as being heavy and awkward to use and wear.

Half-silvered mirrors are also used for in situ visualization. In this method, a standard computer monitor is coupled with a half-silvered mirror or beam splitter glass. The end user looks through the mirror or glass at the real environment and also sees the reflection of the computer display. In the selected publications, only two systems described the use of a silvered mirror as the display device. Peuchot et al. [87] developed a silvered mirror AR system for spinal surgery. Liao et al.’s [63] system uses an IV display coupled with a silvered mirror for MRI-guided interventions.

Lastly, different types of projectors may be used to project the image directly on the patient. Glossop et al. [31, 32] developed a navigation system for craniotomies in which a laser projector projects the planned margins of a craniotomy on the patients’ skull. Tardif et al. [17] examined a projector-based AR system. However, as it was a preliminary work, they made no mention of either the intended application or the type of data to be projected on the patient. Marmulla and colleagues have worked to develop projector-based AR systems [38, 39, 33, 88, 89]: Marmulla et al. [38, 39] developed an AR system for maxillofacial surgery in which a bright projector displays osteotomy lines and tumor margins on the patient. Kahrs et al.’s [33] system for crania-maxillofacial surgery uses two projectors, allowing for the scene to be augmented redundantly or in stereo (for example, with red/green glasses). Volonté et al. [59] developed a projector-based system for laparoscopic surgery wherein volume-rendered color-coded ROIs and segmented objects of interest such as tumors are projected onto the patient.

3.3.3. Interaction Tools

Only a third of the publications specified any interaction tool design concepts or the application of any interaction tools by the end user. The discussion of virtual interaction tools in the selected publications, with few exceptions, is limited to allowing the user to rotate and translate objects [68, 24], to navigate through the virtual scene [72, 43, 22, 24, 44, 23, 57, 35, 18], to use cutting planes [22, 23], to toggle components’ visibility on and off [13, 70, 72, 73, 64, 59] and to change the opacity and color of objects [68, 13, 72, 22, 20, 73, 64].

A few systems mentioned other virtual interaction tools. Wimmer et al.’s [75] HMD AR system allows the user to clip away parts of the anatomy, as well as to specify the size and shape (circle or square) of the actual clipping volume. With Volonté et al.’s [59] laparoscopic system, the user can navigate through the virtual patient’s body (from skin to bone) by changing the window level and width. Furthermore, the user can select particular presets in the color lookup table to highlight information or structures that would otherwise not be visible. Katić et al.’s [13] dental surgery system employs context-aware visualization; the color and visibility of objects depends on the current surgical context, for example, the closeness of a drill to a particular target anatomy. In Bichlmeier et al. [50], the user can specify the thickness of a red border to add emphasis around an ROI of a virtual object. Furthermore, when an HMD is used, the user can change the vision channel to inside the patient, that is, the location of the virtual objects, with simple head movements.

In a number of systems, interaction is necessary to combine the real and virtual worlds. When using a video mixer to combine 3D preoperative models with live video images for neurosurgery, Grimson et al. [68] and Lorensen et al. [69] apply a graphical user interface (GUI) to match laser range data to the 3D models. In the work of Spechtner et al. [6], a radiologist interacts with and controls the AR visualization to support and guide the surgeon. The radiologist views the same scene on a monitor as the surgeon sees in the HMD and registers the preoperative models of the vessel tree to the real-world liver. The radiologist and surgeon are both video- and audio-linked; if different views of the liver are needed to orient the vessel tree, the radiologist can tell the surgeon to reorient the HMD [6]. Zheng et al.’s [40] system incorporates a mouse to correct any segmentation errors before overlaying virtual models onto fluoroscopic images.

A few publications also described the graphical interface of the system. The arthroscopic knee system developed by Tonet et al. [46] uses a four-window interface showing 1) an external viewpoint; 2) a lateral projection; 3) a frontal projection; and 4) an arthroscope tracking light. The lateral and frontal projection windows show the anatomical segments and representation of the axes of the surgical tools, and the arthroscope tracking light window shows both the virtual model and the anatomical area highlighted by the arthroscope field of view. A four-window interface is also used in the endovascular system proposed by Pujol et al. [42]. In their navigation system, three windows provide views of a CT slice (transverse, sagittal and coronal), and the fourth window, giving the virtual view, shows an intraoperative ultrasound image overlaying the endoprosthesis in a 3D mesh model of the aorta. Similarly, in Kosaka et al.’s [49] neurosurgical system, the four-window interface combines a main window that depicts a wireframe mesh object overlaying the mi-
croscope image with three smaller three-slice-view windows (each in a corner) showing the surgical tools on top of CT or MRI slices. The system by Splechtna et al. [6] employs two modes, full screen and split screen, where the active view is outlined with a red frame. In Sauer et al.’s [64] AR thoracoscopic system for spine surgery, the end user can also toggle between two modes, the AR and thoracoscopic views, or use a picture-in-picture view mode. Furthermore, the inset can be shifted either laterally or in depth so that the original image is not obstructed.

In terms of hardware interaction tools, the mouse and the keyboard are most common for data manipulation. The mouse was explicitly mentioned as such in four publications [68, 6, 26, 40]. In Paloc et al.’s [81] mixed reality system for liver surgery, rather than using a mouse, key presses allow the surgeon to switch between monoscopic and stereoscopic views or between real and virtual images.

More sophisticated hardware interaction tools are used in a few works. In Splechtna et al.’s [6] AR system for liver surgery, the radiologist can interact with the system by using both a three-button mouse and a SpaceMouse (a 6-degrees-of-freedom [6DoF] interaction device). The planning software for the microscope-based AR neurosurgery system of Wörn et al. [26] also employs a 6DoF Phantom connected to a virtual tool.

Tangible objects are also used as interaction devices in IGS. Wieczerok et al. [53] implemented a virtual mirror, represented by a laparoscopic camera, to give the surgeon a direct view of occluded objects. In Salb et al.’s [14] system, the end user accesses a virtual graphical interface through the HMD, which is interacted with a Polaris pointer device. Katić et al. [13] allow the end user to interact with the system via a 3D pointer with integrated buttons. Lastly, Lo et al.’s [41] AR cardiac surgery system provides a tangible user interface wherein the user can grasp and physically rotate a heart object to manipulate the virtual object.

Sudra et al.’s [16] endoscopic robotic system employs both speech- and gesture-based interaction. In gesture-based interfaces, some set of motions or configurations of the hands or body are recognized by the system as commands. The surgeon can use speech or gestures to switch between visualization methods, for example, to change parameters or to turn annotation information on and off. Wen et al. [18] are also developing a gesture-based system for their AR needle guidance system for tumor ablation; three hand gestures are recognized: rotation, translation and point selection.

### 3.4. Surgical Domain

All but 12 of the publications [50, 54, 34, 49, 10, 11, 12, 77, 15, 44, 17, 53] specified the type of surgery for which the system is intended. When no surgical domain was specified, the systems described were intended as generic frameworks that can be applied to any surgical domain, or they were simply not described in enough detail. Fig. 1 shows the distribution of surgical domains across the selected publications in a bubble chart.

Most of the papers described navigation systems applied to neurosurgery. This finding is consistent with neurosurgery being the most common and one of the first applications for computer-assisted surgery research. Shuhaiber [90] suggested this frequency may be attributable to neurosurgeons having to resect the smallest possible volumes of tissue in a narrow operative field while trying to minimize damage to the eloquent areas of the brain. Furthermore, in neurosurgery, the surgical anatomy is constrained within a fixed space (the skull) rather than other organs and soft tissues, which allows for feasible registration [90]. A number of the neurosurgery systems are specific to a particular type of surgery, for example, transsphenoidal surgery [21] wherein pituitary tumors are removed through the nose and the sphenoid bone. The MAGI system developed by Edwards et al. [36, 45, 37, 82] is intended for both neurosurgery and otolaryngology, or ENT (ear, nose and throat) surgery. The surgical domain of Lievin and Keenes [27] system includes not only neurosurgery but also cranio- and maxillofacial surgery. The systems described by Glossop et al. [31, 32] and Birklfellner [19] are intended for craniotomy, whereby a bone flap is temporarily removed from the skull to gain access to the brain. The remainder of the publications [68, 83, 84, 69, 3, 29, 4, 10, 30, 71, 26, 91, 28, 20, 92, 72] specified the broader field of neurosurgery as the clinical domain.

Endoscopic surgery, including laparoscopic surgery, is also a common application of mixed reality navigation systems. In endoscopic and laparoscopic surgery, a telescopic rod lens with a video camera is used to guide surgical instruments to the patient anatomy through small incisions or keyholes. Laparoscopic systems were specified for liver [58], digestive [22], abdominal [70], prostate [67], urologic [74] and robotic surgery [73] or simply general laparoscopic surgery [79, 78, 7, 59]. Endoscopic robotic systems were described by Suzuki et al. [73] and Sudra et al. [16]. For the other endoscopic systems [76, 5, 80], the specific type of surgery or surgical task was not specified.

Another common surgical application of the IGS systems is craniofacial and maxillofacial surgery, that is, surgery of the head, neck, face, jaw and other hard and soft tissues in the oral
and maxillofacial regions. Using AR in these types of systems allows osteotomy lines or tumor margins from surgical plans to be transferred to and overlaid on the patient. Furthermore, the visualization of deep structures reduces the invasiveness of this type of surgery [90]. Dental surgery, which also falls under oral and maxillofacial surgery, was described as the surgical domain in two publications [13, 60].

Minimally invasive cardiovascular surgery may also benefit from the application of mixed reality visualization techniques in the OR. A number of publications described the use of mixed reality for different surgical tasks in this domain. A group at the Robarts Research Institute developed an AR visualization system for minimally invasive cardiac procedures, integrating real-time ultrasound and virtual models of the patients beating heart with tracked surgical instruments [43, 61, 41, 8, 9]. Traub et al. [51] also described a system for optimal port placement in minimally invasive cardiovascular surgery.

Liver surgery was cited as the surgical domain in six of the selected publications [58, 6, 65, 81, 22, 57], and breast cancer surgery [66] and breast-conserving surgery [47] were also mentioned. Yet, difficulties arise in developing mixed reality navigation systems for the liver and intestines (pliable organs) or soft tissues such as the breast because they are non-rigid and deform due to heartbeat, breathing, pressure from laparoscopic insufflation, and from being probed [90]. Deformation is not as considerable a problem with surgery of the bone and brain (semirigid organs) [93], which may explain the smaller representation of soft tissue surgery in the field of mixed reality navigation systems.

Orthopaedic surgery, although a common application of computer-assisted interventions, was not as common as an application of mixed reality visualization systems. Five publications specified it as the surgical domain. Tonet et al. [46] described a system for arthroscopic knee surgery; Sauer et al. [64], a thoracoscopic spine surgery system; Peuchot et al. [87], a spinal surgery system; Bichlmeier et al. [54], a vertebroplasty surgical system; and Zheng et al. [40], a system for long bone fracture osteosynthesis.

In general, the systems described were specific to a surgical domain or a particular surgical task. However, a few papers described more generic surgical systems: Samset et al. [48] developed a system for use at the Massachusetts Institute of Technology in liver surgery, cardiac surgery and RF ablation, among others. Liao et al. [63] described a system that can be used in any MRI-guided interventions. Wen et al. [18] described a needle guidance system for RF tumor ablation. Lastly, Lapeer et al. [77] did not specify a particular surgical domain but rather presented a generic framework for using AR in the OR.

That most of the selected publications specified a surgical task or domain suggests that the type of surgery, the steps involved in the surgery and how they are executed motivate the solutions that can or should be used for each component of the mixed reality system.

3.5. Validation and Evaluation

The DVV taxonomy helps to define the core components of a mixed reality system. Both evaluation and validation of these components are essential for demonstrating that a system can and should be used routinely in clinical practice. Whereas validation determines whether the correct system was built, an evaluation of a system shows its value or usefulness.

In terms of data, no publications featured any evaluation or discussion of what types of data would be useful to visualize at particular points in the surgery or for particular tasks. Rather, the selected publications examined only the accuracy of processing methods such as segmentation, reconstruction and registration. Only three papers examined different techniques for visualization processing. Bichlmeier et al. [50] qualitatively examined the visualization of data overlying a cadaver, a phantom and in vivo. Edwards et al. [82] looked at their visualization processing methods in terms of realistic depth perception of the virtual objects. Their results showed an accuracy of about 12 mm of predicting the correct depth of objects with users tending to see objects as deeper than they actually were. Lastly, to examine their focus and context method, Wieczorek et al. [53] had a volunteer wear an HMD while walking around and inspecting a phantom to qualitatively examine the effectiveness of the visualization.

In terms of the view component, a number of papers analyzed the hardware or display technologies that can be used in their system. Wörn et al. [26] and Aschke et al. [4] evaluated different displays in terms of resolution, brightness, contrast, color depth and refresh rate. As a result, they chose to use reflective microdisplays in their system. Linte et al. [43] compared HMDs and monitor displays and determined that the majority of test users found using an HMD more intuitive than a monitor to benefit from the AR experience. Birkfellner et al. [19] compared the use of stereoscopic and monoscopic displays with and without target proximity indication for target localization. Their results showed that stereoscopic vision with target localization had the best target localization success rate (87.5%). Interestingly, after a short evaluation of various displays, Suthau et al. [65] concluded that the currently available AR displays are suboptimal, even HMDs are heavy and awkward and have relatively low resolution.

Although virtual interaction tools were not evaluated or validated in any of the publications, one paper described and validated the system’s hardware interaction tools in terms of the recognition accuracy of the gestures used to interact with the system [13] (recognition rate of over 85%). In addition, Pandy et al. [71] determined the accuracy of the Microscribe arm, which was used for interaction, to be 0.87 mm.

The GUI was evaluated in four publications. Trevisan et al. [24] studied the effect of different interface scenarios (no guidance/visualization, virtual guidance with visualization, augmented guidance with visualization, augmented guidance with visualization and feedback) on the task of cutting into a phantom without touching the dental nerve. Their results showed that the workload of the subject increased with the amount of guidance provided and that the visualization had
more influence on the successful execution of the task than did the actual guidance information. Linte et al. [8] and Peters et al. [9] compared the use of VR and ultrasound image guidance with ultrasound image guidance alone using animal (swine) in vivo experiments. They found that a VR interface improved targeting accuracy and slightly reduced procedure time. For MRI-guided interventions, Liao et al. [63] compared 2D guidance with their IV image overlay system when guiding a needle toward a target. They found that the IV image overlay view reduced procedure time by 75% compared with 2D image guidance alone and Furthermore that the success rate improved by roughly 30%.

Only 32% of the systems were evaluated in terms of their overall use in the OR or the feasibility of bringing the developed system into the OR. A list of the systems that were tested in the OR is given in Table 1. The majority of these results, however, were in terms of qualitative annotations. Surgeons stated that a system was useful [21, 66] or thought that the system provided additional safety [25], and developers found that the confidence of surgeons was improved [82] or simply gained positive feedback from the surgeons [20, 82]. Volonté et al. [59] noted that their system, used to localize tumors under the gall bladder, allowed the surgeon to use the projected preoperative images to adapt their port position, especially in the case of obese patients. Both AV and AR were used in the OR in six patient cases in the work of Paul et al. [72]. The authors found the visualization not only enabled vision beyond the cortical surface of the patient but also gave a larger picture of the surgical area. Both results provided the surgeons with a better understanding of the connection between preoperative patient models and the operative field. Marmulla et al. [38, 39] studied both the use and ergonomics of the system in the OR in two patient cases. They found their navigation system fit within the framework of the surgical procedure and did not prolong the surgery. Furthermore, the surgeon was able to follow the surgical plan with no interference from the navigation tools or system [38, 39]. The orthopaedic system developed by Zheng et al. [40] was used in four patient cases, and the authors stated that the surgical procedure was well supported by their navigation system.

In four papers, the system was evaluated in terms of surgical outcome. Mischkowski et al. [55] found that, after five bimaxillary orthognathic surgeries, the surgeons achieved a precision in the realization of the surgical plan in the range of 1 mm. However, using their X-Scope system significantly prolonged the surgeries (by approximately one hour). Tomikawa et al. [47] evaluated the surgical outcome based on histopathological examinations of the resected tissue of two patients who underwent breast-conserving surgery. They determined that the surgical margins were free of carcinoma cells. The laparoscopic surgical system developed by Teber et al. [74] was used in 10 cases, and in each, definitive histology revealed tumor-free margins. Lastly, Zheng et al. [40] used their orthopedic system in four cases and found no postoperative complications and that all patients healed well.

4. Discussion

This section provides an analysis of the results of our classification with the goal of better understanding the current state of the art in mixed reality visualization systems for IGS. Furthermore, we try to identify the future direction of the field and indicate possible avenues of future research. Where appropriate we reference current research in IGS (beyond that of the selected publications). We begin with some of the limitations of our review methodology.

As mentioned in Section 2, to attain an unbiased selection of papers for review, we conducted a Google query based on particular terms that should or should not appear in the titles. By limiting the search terms so that they had to appear to be within the title, using the google term allintitle, (rather than anywhere in the article), we were able to decrease the initial database from over 100,000 articles to just under 500. This reduction, however, came at the cost of certain papers not being identified with the specified query, for example, Stetten et al.’s work on merging ultrasound images with direct vision [94, 95]; Jannin et al.’s AR system in which contours from functional information are overlaid onto the patient by means of the microscope [96]; and Ghavaghan et al.’s portable image overlay device for laser projection of preoperative models on the patient during open liver surgery [97]. Although each of these papers meets our inclusion criteria in that they define mixed reality systems developed for IGS, they did not appear in our search results because their titles did not feature the terms reality or virtuality. Although ours is not an exhaustive list of publications in the field, we believe that both the results of our analysis and the trends we identified and discuss in the following section remain valid.

4.1. Data

Patient-specific imaging data were well defined in the selected publications, with different imaging modalities used in different surgical domains. With the development of new imaging modalities, it is becoming increasingly important to focus on which of the available data formats should be shown to the end user at a given surgical step and to study methods of combining the different data sources into a coherent visualization. Showing too much or all of the available data is often not desirable as it may confound the viewer and make it difficult to focus on particular features or information. This problem may become somewhat resolved as imaging solutions become more targeted to given surgical tasks for surgical situation awareness systems.

4.1.1. Analyzed Imaging Data

Although faster processors and graphics cards have allowed for real-time frame rates of overlay images in mixed reality visualizations, the standard analyzed data representation of patient anatomy remains a surface rendering. The systems in 42 papers used surfaces either alone or in combination with another data type, and 11 papers used wireframe data representations (Fig. 2). That the majority of the selected publications used simple data representations (e.g., points, lines, contours,
The main disadvantage of surface (or wireframe) rendering is the need to pre-process the data in order to delineate all structures of interest. Manual segmentation is a long and time-consuming process, and although automatic segmentation methods exist, variations in image quality mean the results of these methods cannot be guaranteed in clinical routine. For this reason, an expert may be required to check the results of the segmentation [98].

Volume rendering was used alone or in combination with another analyzed data representation in only 11 of the selected papers. The main advantage of volume rendering over surface rendering is that there is no need for delineating the object of interest. However, the lack of delineation makes it more difficult to visualize objects of interest individually; to compute the volume of objects of interest; and to simulate resections of particular structures without cutting neighboring structures [98]. The need to find appropriate transfer functions to properly display and show structures that should be delineated may be complex and time consuming.

Recently, Kainz et al. [99] investigated the absence of 3D volume visualizations in clinical practice by surveying 24 experts on the topic. The results of their study suggest that, although physicians prefer 2D multiplanar reconstructions of 3D data to 3D visualization methods for interventional planning, there is a need for 3D visualizations and they are welcome in the clinical domain. This need, however, is specific to the type of diagnostic and interventional application [99]. The results of this study further emphasize the need to validate visualization processing methods, the second component of the DVV taxonomy. By applying proper metrics to validate new visualization methods and compare them with currently used ones, the field can determine whether there is a need for new methods and how these methods can aid in both the surgeons decision-making process and the execution of particular surgical tasks.

### 4.1.2. Prior Knowledge Data

Surprisingly, prior knowledge data and derived data were not often reported in the selected publications. In terms of prior knowledge data, papers described the depiction of tracked surgical tools overlaying the preoperative patient models. Tools were depicted as simple wireframe or surface models, typically as simple cylinder or line renderings. However, such simplistic depiction techniques are problematic because of the difficulty of locating tools at the proper depth within the patient anatomy. As Bichlmeier et al. [50] suggested, more careful consideration is needed to determine how to represent surgical tools so as to improve, rather than confuse, localization and navigation. Future work in this area may include exploring lighting effects such as shadow-casting, using more sophisticated occlusion techniques and depicting more photorealistic tools.
ference anatomies and labels, were mentioned in the selected publications. Yet, using prior knowledge data is neither infeasible nor novel. For image-guided neurosurgery, Nowinski et al. [100] used a brain atlas database to support visualization of anatomical structures for real-time navigation. St-Jean et al. [101] visualized patient-specific MRI data registered to a deformable volumetric atlas of the basal ganglia and thalamus for planning and guidance in neurosurgery. Clements et al. [102] used preoperatively computed liver atlases to deform intraoperative images of the liver. Atlases and reference models continue to be used for preoperative planning; however, they do not seem to be the focus of current research in IGS, in particular, in mixed reality environments and visualization techniques. This is perhaps due to the current focus on more targeted and patient-specific visualization solutions.

4.1.3. Derived Data

Derived data was presented mostly in terms of direct measurements such as distances, clinical scores and derived tool trajectories. This type of data was typically presented on screen using simple analyzed data objects such as lines and points. A future area of research may examine how to effectively visualize the fusion of derived data such as clinical scores with that of visualized imaging data such as MRIs.

In the selected publications, only Linte et al. [43] applied uncertainty data in the form of 95% confidence intervals based on TRE scores. Other researchers, however, have also looked at conveying uncertainty measurements to the end user: Risholm et al. [103] computed registration uncertainty and visualized it in terms of color-coded isocontours. Simpson et al. [104, 105] visualized registration uncertainty using an uncertainty cone that represented the distribution of a possible planned linear path. Najafi et al. [106] depicted the uncertainty of tracking sensors using ellipsoids. Depicting uncertainty data has been shown to help in localizing targets [104] and in improving registration accuracy [106]. For the most part, however, the depiction of uncertainty data has remained academic. To our knowledge, only the Medtronic cranial application (Medtronic, USA), provides any uncertainty visualization, in terms of TRE scores displayed using isocontours.

Given that a surgeon relies on visualization data for decision making and navigation, visualizing uncertainties that arise from registration, segmentation and tracking provides important information that may affect surgical decisions. In the particular field of mixed reality visualization for guidance and navigation, uncertainty information may be especially useful for visually depicting the accuracy of the registration of the real and virtual objects. Further clinical experiments on what type of uncertainty information is needed and how it should be visualized are necessary to better understand when uncertainty is beneficial to a surgeon.

4.2. Visualization Processing

In over one-third of the papers there is no mention of what techniques were used to transform the data to present it visually to the end user (Fig. 3). In the remainder of the papers, simplistic visualization processing techniques were used, the most popular of which were color-coding structures of interest and using transparency to merge different modalities of data. Unfortunately, such simplistic techniques often confound the understanding of structures and their spatial relationships. It has been shown, for example, that using transparency to merge imaging data can confuse the perception of relative depths and spatial relationships between surfaces [107]. Furthermore, when transparency is combined with stereoscopic viewing, the perception of the depth of the stereo images becomes ambiguous [108]. It is surprising then that more sophisticated techniques have not been more thoroughly explored for a more comprehensive understanding of, and interaction with, complex medical multimodal data in an intraoperative environment.

4.3. View

The view component of the DVV taxonomy was generally well defined, with the exception of the interface classes.

4.3.1. Display

To assess the use of the different display technologies applied in mixed reality IGS systems, we examined the types of displays used across all of the selected publications, as well as those in each surgical domain for which a system was described.
in more than 10 papers (see Fig. 4). For surgical domains mentioned in fewer than 10 papers, it was not possible to identify any significant trends.

The most common display device across all surgical domains was a monitor (used in 47% of the selected papers). The prevalent use of a monitor as the display device may result from its current presence in the OR, which may ease the transition from a navigation system to a mixed reality IGS system. Further advantages of using a monitor as display device include the ability of multiple users benefiting from the visualization and a greater range of usable visualization techniques. Because it is not necessary to introduce a new display device into the OR, using a monitor is both cost effective and non-intrusive. Monitors were the most common display device for endoscopic and laparoscopic systems. The use of a monitor for traditional endoscopic and laparoscopic surgery creates a natural platform for AR techniques wherein virtual models are merged with the real-time video images from the endoscope.

For neurosurgery, the microscope was the most common display device (52%); this finding reaffirms that a display device already present in the OR may be a logical solution for visualization. Applying the microscope to show a AR visualization requires no additional display device and little to no additional cost and also reduces disruption to the surgeon’s workflow.

For maxillofacial surgery, the most common device was the projector (38%), with HMD, monitor and portable screen having roughly similar distributions. Such a spread may result from the lines or contours (e.g., osteotomy lines, tumor contours) in plans projected on the patient surface being sufficient to guide surgeons in craniofacial, maxillofacial and dental surgeries.

None of the selected publications identified holography as a display device. However, Liévin and Keene [27] mentioned that they are experimenting with using holography rather than a monitor. Holography has been presented as a solution for mixed reality visualization by a number of research groups. Ko et al. [110] studied the use of projecting 3D holographic images on the surgical site to serve as a visual template through which surgeons can view the patient during surgery. Bergman et al. [111] used digital holography in neurosurgery, in particular for cyst cannulation. They employed transparent holograms of a patient’s ventricles, vessels, sulci and other brain structures to visualize the spatial relationships between structures, ultimately aiding the surgeon in finding the optimal needle trajectory to the target cyst. Nowatzyk et al. [112] presented initial work toward replacing their previous AR system, in which ultrasound slices overlay the patient via a half-silvered mirror [95], with real-time holographic tomography. More recently, Sittler et al. [113] worked on developing a holographic combiner (i.e., a beam splitter that redirects the projected image such that the field of view and projected infinity image can be seen simultaneously) for use in HMD AR surgical systems. The advantage of holographic displays is that holograms are highly realistic and allow viewing from a wide range of angles [114]. However, the current limitations of holographic devices (including allowable image size and recording media that do not allow the image to be easily refreshed) still limit the use of this technology in the OR.

Mobile devices, such as the iPod touch and iPad are also finding their way into the OR. One example of this is Smith & Nephew Dash’s portable navigation system powered by Brainlab. This orthopaedic navigation system guides the surgeon to accurately place knee and hip implants. Such ubiquitous devices, which are small and portable, are increasingly being used across many medical domains and may soon become standard display devices in the OR.

Combined with the trend of preferring particular devices in a specific surgical domain, the variety of display devices used across the various surgical domains suggests that there is no one global IGS solution. Analysis of the surgical domain, the surgical tasks, the surgeons’ workflow and the OR environment should help define the requirements for the appropriate display device.

4.3.2. Perception Location

Perception location was examined across all publications as well as for each surgical domain for which more than 10 papers described a system for that domain (see Fig. 5). Across all selected publications, patient and monitor were the perception location most frequently used, with a slight preference for the patient (48% of papers used the patient, and 44% used a monitor). In a few papers (3%), the perception location was not specified, and in a few others (5%), the end user could view the visualization on either the patient or a monitor.

In papers that dealt with neurosurgery and maxillofacial surgery, the patient was the most often used perception location (64% of the time). As with the selection of the display device, this frequency suggests the importance of not disrupting the surgeon’s workflow. Using the patient as the perception location ensures that the surgeon need not look away from the surgical scene to benefit the visualization. This choice also reduces the burden of having to transform guidance images from the navigation system to the patient on the table, a task that is non-trivial especially when a surgical microscope is being used as in neurosurgery [45].

For endoscopic and laparoscopic navigation systems, the monitor was the most the common perception location (57%). As a monitor is traditionally employed in such surgeries, its use as the perception location allows the easy extension of traditional navigation tools and requires no training on the surgeon’s part. Still, the patient was cited as the perception location in almost one-third of the selected papers. This rate may suggest a move toward using the patient as perception location in most surgeries, even those wherein the surgeon traditionally does not look directly at the patient.

Although AR researchers claim there is a need to minimize disruption to workflow during surgery and therefore largely choose the patient as perception location, to our knowledge, no study has looked at how different perception locations affect the surgeon, their workflow or the surgical outcome. It is conceivable that occasionally looking away from the surgical scene may allow for a reassessment of the current surgical situation or even a short mental break. It is also possible that having multiple perception locations, with differing locations used at
selected steps throughout the surgery, may be a desirable solution. Ergonomic and interface studies that examine the use of multiple perception locations are needed to better understand the effects of choosing one location over another at a particular surgical step.

4.3.3. Interaction Tools

Interaction tools were described in less than half of the selected publications (Fig. 6). Only 18% of the papers mentioned virtual interaction tools, in terms of how the user can manipulate the view and visualization of virtual objects, and only 16% described the hardware interaction tools (e.g., mouse, keyboard, pedal, tangible objects, etc.) that can be used to interact with the system. Furthermore, only 7% of papers described the look of the GUI.

The most common interaction tools remain the mouse and keyboard. Such classical interaction hardware paradigms, however, are not appropriate to solving interactions between surgeons and visually processed data in augmented environments [115, 116]. Furthermore, because of the need for sterility, it is typically a technician or an assistant who works with the mouse or keyboard upon receiving verbal instructions from the surgeon. Such a verbal communication interaction paradigm is often slow and prone to errors and misunderstandings resulting from verbal ambiguities [117]. To overcome the limitations of such interactions, a number of solutions have been proposed.

A few of the selected papers described novel hardware interaction devices: the virtual mirror developed by Bichlmeier et al. [115, 116], the SpaceMouse used by Splechna et al. [6], the Microscribe by Pandya et al. [28], the tangible heart object by Lo et al. [41], the Polaris pointer by Salb et al. [14] and the gesture-based interactions by Sudra et al. [16] and Fischer et al. [118]. Gesture-based interactions were also recently presented by Schwarz et al. [119]; in their system, users wearing inertial sensors train the system to recognize particular gestures for desired interaction tasks. Such gesture-based interaction methods may better fit a surgeon’s workflow because they allow both gesture customization and easy consideration of the different constraints on the surgeon in the OR.

Onceanu and Stewart [117] proposed a different type of solution based on using a tool already present in the OR. Their device, which consists of a base into which a tracked surgical probe is placed, allows joystick-like interaction such that the az-
imuth and elevation angles determine the speed and direction of an on-screen cursor. Clicking is possible by pressing the probe further into the base. In a user study, the authors compared the input device with a mouse and verbal communication. Their results showed that, although faster than both dictation and the joystick, the mouse was not significantly more accurate. That users found the mouse significantly faster is not unusual, as it is a traditional and familiar device. Furthermore, their results suggest that metrics other than accuracy and speed are needed to enable a comparison of devices and methods. An important avenue of future research will be additional work on testing and developing hardware devices that are more appropriate to interactions with visually processed data and to easily and efficiently adjusting viewing and visualization parameters.

Another possibility for future work in interaction tools is to create interfaces that require little to no interaction. In the domain of surgical interventions, requiring the surgeon to manipulate data using particular tools may interfere with and interrupt the surgical workflow. It is possible that, in an ideal IGS system, the view and visualization parameters would change such that a suitable representation of the appropriate data would be presented automatically at any given stage of the surgery. Research in this area was recently presented by Katić et al. [13], who developed a context-aware visualization method. In their system, the visualization of virtual objects, including their color and visibility, changes depending on the particular surgical context. Other researchers have looked at presenting only the necessary data and information at a particular surgical step either by determining a surgical model [120], by monitoring the signals and actions performed in the OR [116, 112] or by classifying microscope video images to automatically recognize the current surgical phase [122, 123]. In such systems, the surgeon sees the most appropriate aspects and views of the data, allowing for fewer manipulations of the data and therefore fewer disruptions to the surgical task at hand.

The GUI was only discussed in 7% of the selected publications. Typical GUIs for navigation remain three- or four-window views, with three windows showing slice views and a possible fourth window showing a volume rendered or augmented view. Although a few papers [24, 9, 8] looked at evaluating the user interface, typically in terms of comparing a mixed reality interface with a traditional navigation interface, this remains an under-represented topic in the literature.

4.4. Validations and Evaluations

System evaluations were performed in 87% of the selected publications, yet few papers looked at evaluating or validating the DVV components (see Fig. 7), and no publications examined all of the components.

The majority of the researchers (36%) looked at validating the systems in terms of either the accuracy of the system as a whole, or the registration, calibration or overlay accuracy of the real and virtual images (see Fig. 8). In general, although some system components were validated using numerical methods or phantoms, the majority were not evaluated in clinical settings on real patients.

In the OR, the IGS system, the surgeon and the patient interact in a three-way relationship: The surgeon operates on the patient and interacts with the IGS system through a human-computer interface, and the IGS system may provide intraoperative imaging of the patient. Based on these components and their interrelationships, Jannin and Korb [9] proposed to classify possible criteria for assessing IGS systems into six families: criteria related to the patient, the surgeon, the IGS system, and the interactions between surgeon and patient, IGS system and surgeon, and IGS system and patient. In addition to these criteria, Jannin and Korb [9] identified six levels of assessment based on evaluating different properties of the IGS system. Levels one to six comprise the systems technical parameters, its reliability in a clinical setting, its efficacy in terms of surgical performance, its effectiveness in terms of patient outcome, the economic aspects of its use and, lastly, the social, legal and ethical aspects of its use.

The specifications of these criteria and assessment levels outline the complexity and difficulty involved in assessing IGS surgery systems. Assessing patient-related criteria in terms of surgical outcomes such as cosmetic results, pain and clinical scores is particularly challenging and time consuming, for example, in planning and carrying out clinical trials and measuring improvement based on whether a particular technology or technique was used. This limitation was made evident by the lack of systems in our selected publications that were used in a real clinical setting.

The process of evaluating IGS systems is complicated by the fact that current methods and metrics have been optimized for existing solutions. It is important to go beyond validation studies based on accuracy alone and consider all of the criteria involved in using an IGS system. New metrics for assessing the
different criteria need to be developed, and these require a focus on working with surgeons and physicians to define metrics that can be used not only to examine systems individually but also to compare their components and even the systems themselves.

5. Conclusions

This paper provided a review of mixed reality IGS systems and an analytical perspective of the recent solutions that have been explored across different surgical domains. Such a review may aid developers in exploring current solutions and developing novel solutions to problems that have yet to be solved and therefore, may help in bringing new systems into routine clinical use.

The DVV taxonomy facilitated our review. We believe that the elements of the DVV taxonomy should be given in a publication describing such any mixed reality visualization system in order to give a good overview of the system and allow for evaluation of their systems as well as across systems. In Fig. 9 we provide a check-list based on the DVV taxonomy that can be employed when developing and describing a new or existing mixed reality guidance system.

![Check-list based on the DVV taxonomy. The three main factors of the taxonomy (Data, Visualization Processing, View) are given with their subclasses and common instances in brackets.](image)

In surveying the current landscape of mixed reality IGS systems, we identified a lack of focus in 1) choosing appropriate data to be visualized for particular surgical scenarios; 2) applying visualization processing techniques; 3) proposing interface solutions; and 4) evaluating and validating systems and their individual components. In analyzing the chosen solutions for the DVV components of the mixed IGS systems in the selected publications, we identified certain trends; however, the validations and evaluations of these trends are lacking. Furthermore, the solutions presented are often based on available technology, anecdotes and incomplete knowledge.

In terms of system evaluation, entire systems and their individual components must be validated and evaluated to show that they are usable not only in ideal circumstances, but also under the worst conditions when the surgeon is fatigued and when clinical complications occur. Our review showed that psychophysical and human factor studies of visualization methods and interaction techniques are particularly lacking and should become a focus of future work. Engineers and computer scientists must work closely with clinicians to develop new metrics that can be used consistently to examine the effectiveness of particular solutions such that the same criteria is used to compare solutions across different systems. Both the development of systems that meet the surgeon’s requirements and the evaluation and validation of each system component will be of great importance for the introduction of more mixed reality systems into the OR.

The variety of the publications reviewed demonstrates that there are many different technological solutions and tools for each of the DVV components. The focus in the field, therefore, must now turn from technical innovations to methods of combining these tools to develop systems that fit seamlessly into the OR and aid the surgeon in specific tasks. System developers must work closely with clinicians in order to determine the surgeons requirements in the OR. To achieve this goal, targeted user studies should be done that aim at learning the needs of surgeons and the constraints of the complex OR environment.

Acknowledgments

The authors would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC), Canadian Institutes of Health Research (CIHR MOP 97820, MOP-74725 & CHRP), and INRIA’s “Programme Sabbatiques”.

References


[31] C. Bichlmeir, T. Selhorst, S. M. Heinig, N. Navab, Improving depth perception in medical ar: A virtual vision panel to the inside of the pa-


