

Using a depth-sensing infrared camera system to access and manipulate medical imaging from within the sterile operating field

Matt Strickland, MD, BASc*

Jamie Tremaine, BASc[†]

Greg Brigley, BASc[‡]

Calvin Law, MD, MPH[†]

From the *Division of General Surgery, Department of Surgery, University of Toronto, The Banting Institute; the [†]Division of General Surgery, Sunnybrook Health Sciences Centre and the University of Toronto, Toronto, Ont. [‡]Independent researchers, Toronto, Ont.

Correspondence to:

M. Strickland
Division of General Surgery
Department of Surgery
University of Toronto
The Banting Institute
100 College St.
Toronto ON M5G 1L5
matt.strickland@utoronto.ca

DOI: 10.1503/cjs.035311

Background: As surgical procedures become increasingly dependent on equipment and imaging, the need for sterile members of the surgical team to have unimpeded access to the nonsterile technology in their operating room (OR) is of growing importance. To our knowledge, our team is the first to use an inexpensive infrared depth-sensing camera (a component of the Microsoft Kinect) and software developed in-house to give surgeons a touchless, gestural interface with which to navigate their picture archiving and communication systems intraoperatively.

Methods: The system was designed and developed with feedback from surgeons and OR personnel and with consideration of the principles of aseptic technique and gestural controls in mind. Simulation was used for basic validation before trialing in a pilot series of 6 hepatobiliary-pancreatic surgeries.

Results: The interface was used extensively in 2 laparoscopic and 4 open procedures. Surgeons primarily used the system for anatomic correlation, real-time comparison of intraoperative ultrasound with preoperative computed tomography and magnetic resonance imaging scans and for teaching residents and fellows.

Conclusion: The system worked well in a wide range of lighting conditions and procedures. It led to a perceived increase in the use of intraoperative image consultation. Further research should be focused on investigating the usefulness of touchless gestural interfaces in different types of surgical procedures and its effects on operative time.

Contexte : À mesure que les techniques chirurgicales dépendent de plus en plus des appareils et de l'imagerie, il devient plus important que des membres « stériles » de l'équipe chirurgicale aient libre accès à la technologie non stérile dans leur bloc opératoire. À notre connaissance, notre équipe est la première à utiliser une caméra infrarouge peu coûteuse dotée de capteurs de profondeur (une composante du Kinect de Microsoft) et un logiciel maison pour donner aux chirurgiens une interface gestuelle sans contact avec laquelle ils naviguent dans leurs systèmes d'archivage et de communication peropératoires.

Méthodes : Le système a été conçu et développé en collaboration avec des chirurgiens et du personnel du bloc opératoire et tient compte des principes de l'asepsie et des contrôles gestuels. La validation de base a reposé sur des exercices de simulation avant le lancement d'une série d'essais pilotes au cours de 6 interventions hépatobilio-pancréatiques.

Résultats : L'interface a beaucoup servi lors des 2 interventions laparoscopiques et des 4 interventions ouvertes. Les chirurgiens ont principalement utilisé le système pour établir des corrélations anatomiques et des comparaisons en temps réel entre les échographies peropératoires et les tomographies et imageries par résonance magnétique préopératoires et pour faire de l'enseignement aux résidents et aux stagiaires en formation postdoctorale.

Conclusion : Le système a bien fonctionné dans des conditions de luminosité et lors d'interventions diverses. Il a donné lieu à plus grand recours perçu à la consultation des images peropératoires. Des recherches plus approfondies devraient porter sur l'utilité des interfaces gestuelles sans contact dans différents types de chirurgies et sur leurs effets sur la durée des interventions.

Technological systems supporting the surgical team play an increasingly important role in hospitals and health care. It is commonplace for operating rooms (ORs) to be outfitted with computers that allow access to picture archiving and communications systems (PACS), patient records via the hospital's electronic medical record (EMR) software, computerized physician

order entry systems and OR management software suites.

Unfortunately, the necessary divide between the sterile operative area and the nonsterile surrounding room means that, despite physical proximity to powerful information tools, those scrubbed in the OR are unable to take advantage of those resources via traditional human-computer interfaces. In the case of medical imaging, surgeons resort to studying the case beforehand, asking circulating nurses or onlookers to control the devices for them, or using ad hoc barriers (e.g., a sterile green towel over the mouse) to navigate to the necessary information. Alternatively, a surgeon may choose to control the equipment himself, but this results in his contamination and the necessity for time-consuming rescrubbing and an incentive to minimize his reconsultation of the available imagery. In current practice, therefore, the use of modern technology in the OR is at best awkward and fails to realize its full potential for contributing to the best possible surgical outcomes.

Early attempts to overcome the OR sterility barrier were mainly based on speech recognition, and some systems were even commercialized as part of integrated OR suites.^{1,2} In one such system, the surgeon wears a microphone headset and wireless transmission unit at his waist throughout the procedure. He can activate the system using a keyword and then manipulate settings by giving simple voice commands, such as “insufflator...pressure...down.”³ Although the use of speech recognition in the typically noisy OR environment has been studied and shown to be feasible,⁴ the approach has several fundamental drawbacks.⁵ While voice commands are appropriate for simple Boolean functions like turning devices on or off, they become cumbersome for more complex tasks, such as moving a cursor in 2 dimensions or navigating a typical EMR.

Advances in machine vision have led several groups to develop noncontact gestural systems for the operative setting. Graetzel and colleagues⁶ first reported a system that used computer vision to replace standard mouse commands with hand gestures. They were able to follow the surgeon's hands using a stereoscopic colour camera arrangement and tracking algorithms that relied on both colour and processed depth. Other groups have since reported similar results.⁷ More recently, Wachs and colleagues⁸⁻¹⁰ developed another gesture-based system using a simple webcam for data capture. Without gathering depth information, their group was able to devise a series of gestures to operate a PACS system and manipulate images intraoperatively.^{8,9} The system was used successfully in a neurosurgical biopsy procedure in 2008.¹⁰

Stereotactic, image-guided surgery has long used image registration and proprietary pointing devices to allow the surgical team to interact with the system.^{11,12} One group has reported moderate success tracking the surgeon's face to help control a laparoscope.¹³ Solutions to a related but independent problem, using gestural interfaces to improve radiologists' ability to navigate imaging, have also been

proposed using commercially available, nonsterilizable hardware, such as the 5DT DataGlove¹⁴ and a controller for the Nintendo Wii.¹⁵

We saw the possibility of another approach with the 2010 introduction of the Xbox Kinect by Microsoft. The Kinect is an after-market add-on to the company's popular Xbox gaming system that allows users to control video games without a physical controller. The device uses an infrared (IR) depth-sensing camera system to track user movements, which software then translates directly to motion or actions on the television screen.¹⁶ Shortly after its release, enthusiasts decoded the Kinect's USB data stream, and low-level drivers were released publicly to spur creativity and novel applications of the technology.¹⁷ We sought to develop a suitable prototype design and conduct a series of pilot procedures to examine the feasibility of this new class of input device to help bridge the OR sterility barrier and eliminate the time and space gap that currently exists between imaging review and visual correlation with real-time operative field anatomy. We report our findings here.

METHODS

We conducted this study in 3 major phases: system design based on predetermined specifications, creation of a gestural interface in a simulated setting and, finally, pilot-testing in a set of live OR procedures.

Specifications

The system was initially designed and created outside of the OR in a closed laboratory. This phase focused on developing the appropriate hardware, software and user interface. Existing literature and informal interviews with surgeons, residents and nursing staff helped establish the following list of specifications.

1. The system must allow for noncontact control of a PACS computer, the IT system most likely to influence intraoperative decision-making.
2. To ignore inadvertent commands, the gesture recognition system should be inactive until hailed by a distinctive action. The system should be locked using another distinctive action.
3. The system should use information from the operator's upper limbs and torso to implement the basic functionality of a mouse-like device.
4. Gesture recognition must be robust and reliable.
5. The user interface must have minimal equipment requirements and account for user fatigue and unintentionality while optimizing intuitiveness, real-time interaction and ease of learning.¹⁸
6. All gestures must abide by the constraints imposed by OR rules for aseptic technique and working in close quarters with assistants.
7. The system should be easy to integrate into existing ORs

with minimal distraction, training or human resources.⁶

We accepted further constraints in our design, as suggested in the Association of Perioperative Registered Nurses guidelines for working practices in an OR.¹⁹ These basic rules help define a 3-dimensional area in which gestures should ideally be performed. This zone extends roughly from the waist inferiorly to the shoulders superiorly and from the chest to the limit of the outstretched arms anteriorly and to about 20 cm outside of each shoulder laterally. For the initial release, gestures were implemented to emulate scrolling with a scroll wheel, cursor navigation in 2 dimensions and full mouse button functionality. We included additional gestures to calibrate the system, lock it and unlock it. These latter gestures are essential to ensure no commands are sent inadvertently while the team is operating — a requirement discussed by previous investigators.^{6,20}

Pilot clinical study

Once we completed the final prototype design, testing was planned in cooperation with a single surgeon (C.L.). Both laparoscopic and open procedures were targeted. We obtained feedback after each procedure to gain insight into the system's use and its perceived strengths and weaknesses.

RESULTS

System overview and design

The final system design was fully integrated onto a portable cart, including an IR camera unit, image-processing unit, feedback display, PACS-equipped OR computer and PACS

system display. The cart-based approach helps minimize set-up time and can easily be moved anywhere in the room to accommodate surgeon preferences and procedure requirements.

For the IR camera system, we used a Microsoft Kinect. This hardware device is generally used as an input device for Microsoft's Xbox gaming console and includes an IR depth-sensing camera, an RGB camera and a series of microphones. In this application, the Kinect provides a low-cost depth sensor from which we can extract 3-dimensional scene information. An alternative device, the Asus Xtion, offers the same relevant functionality for about the same price. The scene data from the IR camera is sent to an image-processing unit that interprets it and extracts information about the user's position and motion. The processing unit offers visual feedback to the user with regards to their actions and calibration status via a second monitor. Based on the user's gestures, the processing unit sends an output signal to the standard, PACS-equipped OR computer.

The PACS-equipped computer was similar to our facility's pre-existing OR computers. It was connected to the hospital's internal network and equipped with licensed software to display imaging and radiology reports. Before the procedure, the surgeon is able to load the images required for the operation, as per usual practice. The video output from this computer is then displayed on a PACS monitor.

This entire system is illustrated schematically in Figure 1 and photographically in Figure 2.

Usability tests

Preliminary testing occurred in an unused OR to ensure

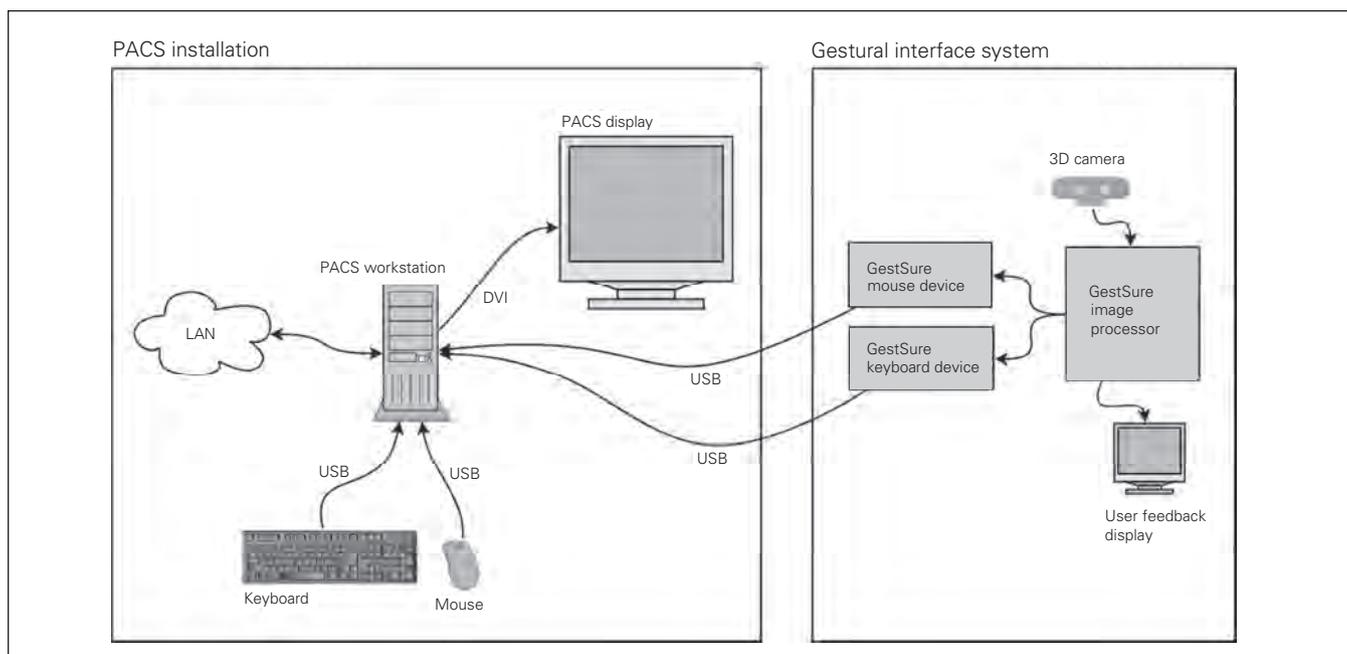


Fig. 1. System topology. PACS = picture archiving and communications systems.

the area was suitable and that lighting, surgical workflow and any potential sources of interference were accounted for. Four test users, including 1 surgeon (C.L.) and 1 surgical resident (M.S.), were introduced to the system and asked to perform a series of tasks meant to simulate typical OR situations. The users were charged with accessing different forms of imaging (computed tomography [CT], magnetic resonance imaging [MRI], plain film), to do so

under different operating configurations (mock patient in decubitus, lithotomy and supine positions), and with the system situated at different distances and locations in the room. We obtained feedback via user interviews and through observation of their performance during the tasks. The testing revealed that the gesture library was generally intuitive and easy to learn, meeting the previously described specifications. Testers quickly became comfortable with the system, and all were able to master the simple set of gestures within 10 minutes. The participants' strong recall of the gestures during subsequent testing supported the intuitive nature of the system.



Fig. 2. Gestural interface system.

Pilot study

We included a total of 6 procedures (2 laparoscopic, 4 open) in the pilot study. With the exception of the patient in procedure 6 (details follow), all patients were discharged in the expected time frames, without major complications.

Procedure 1: laparoscopic adrenalectomy in a 59-year-old man with adrenal adenoma and Conn syndrome

In this laparoscopic adrenalectomy, the system was accessed sparingly and only for basic anatomic correlation. However, the system performed reliably, especially in the dark, laparoscopic environment.

Procedure 2: laparoscopic pancreatectomy in a 67-year-old woman with pancreatic intraductal papillary mucinous neoplasm

In this spleen-preserving laparoscopic pancreatectomy, the system was accessed frequently to determine the trajectory of the splenic vasculature. The most used gesture in this series was scrolling animation of axial series in caudad and cranial directions. The system also performed well in the dark conditions and was felt to substantially enhance the procedure because it was easy to use and removed the need to break sterility to access imaging.

Procedure 3: hepatic resection in a 67-year-old man with diffuse metastatic neuroendocrine tumours

In this extended hepatic resection, a total of 25 lesions were excised, guided by the use of intraoperative ultrasound (IOUS). It was noted that access to the CT/MRI scans using the touchless interface allowed for real-time correlation of IOUS to CT/MRI and facilitated faster lesion targeting owing to this ability.

Procedure 4: hepatic resection in a 48-year-old man with metastatic colorectal cancer

This and the preceding hepatic resection were both considered more complex than a standard hepatic resection owing to the need to target small lesions for parenchyma-preserving resection techniques. Access to the imaging was used

frequently during these hepatic resections, even outside of IOUS, for target and anatomic structure acquisition.

Procedure 5: Whipple procedure in a 74-year-old woman with pancreatic adenocarcinoma

In this Whipple procedure with portal vein resection and reconstruction, a vein resection was anticipated before the procedure. It was deemed necessary during the case, and it was felt that this decision was facilitated by ready access to the imaging system using the touchless interface. Access to imaging was used frequently during this operation and was used to target margins for the venous resection. The final pathology confirmed a margin-negative venous resection of the pancreatic adenocarcinoma. It was perceived by the surgeon that imaging was used more frequently in this procedure than in his prior experience.

Procedure 6: palliative resection in a 43-year-old man with recurrent renal cell carcinoma

This was a complex procedure for palliative resection of a recurrent renal cell carcinoma that had caused a perforation of the duodenum and retroperitoneal sepsis. A complex distal gastrectomy, duodenal exclusion and duodenal fistula repair was undertaken. The imaging system was used frequently throughout the case to help localize anatomic structures in real time during the procedure. It was felt to be of assistance to the conduct of the procedure. The patient, however, had further septic complications during his hospital stay but eventually returned home safely. It was felt that these complications were related to the nature of his disease.

The most commonly used gestures in the open series of procedures involved scrolling animation but also series selection. Windowing was not commonly accessed, and it was felt to be owing to the presence of multiphasic scans, which made series selection a more vital gesture.

Besides anatomic correlation, the system was also used frequently for intraoperative teaching of residents and fellows, mainly to help illustrate approaches and potential hazards.

Challenges

Issues that arose during the use of the touchless system mostly related to capturing the surgeon's gestures for system activation in a crowded operative space. Some of the challenges may have been related to the heat from the overhead halogen OR lights, a known source of interference for this type of IR depth-sensing camera.

DISCUSSION

Another group has recently reported using similar hardware to extract data for gestural interpretation for possible use in surgical settings,²¹ but to the best of our knowledge,

the work we describe here is the first time an IR depth-sensing camera has been used intraoperatively as part of a human-machine interface. Furthermore, it is the first time a vision-based system has been used so extensively for image navigation outside of stereotactic surgery.

The type of camera used, most readily available as a component of Microsoft's Kinect controller, is now low-cost and well suited for the OR environment. It gathers rich, high-quality 3-dimensional information that can be reliably parsed to extract gestural information. It works well in a wide range of lighting environments, including complete darkness, making it suitable for both open and laparoscopic procedures.

Because the hardware relies on projecting structured IR light on the scene and then retrieving depth information based on how this known pattern is distorted, it is vulnerable to interference by other sources of IR light. The most notable source in many ORs is the overhead lights used to illuminate the surgical field. In the operating suites where we tested the system, the traditional halogen lights were found to interfere substantially with the camera's ability to acquire 3-dimensional data when the user was directly illuminated. This weakened the robustness of the system when the user tried to use gestures directly over the incision where, generally, the lights are focused. This forced the surgeon to turn away or step back from the operative field to control the system. This problem should be allayed further by the increasing use of low-IR, light-emitting diode lighting in ORs.

Some authors have reported that Medsonic's StealthStation, a computer-assisted surgery system that uses optical IR tracking, can interfere with pulse oximeters used for patient monitoring.^{22,23} Throughout our pilot study, there were no abnormalities reported in the pulse oximetry readings that could be attributed to the depth-sensing camera's IR projection. Based on the IR wavelength projected (830 nm) and the diffuse, low-power illumination pattern,²⁴ we would expect its likelihood of interference with pulse oximeters to be low.

Feedback gathered from surgical staff in the pilot hepatobiliary-pancreatic procedures indicated that the most useful functionality was the ability to intraoperatively animate CT and MRI scans and switch to different series within a study. Abilities that ranked very low in importance included windowing, zooming, rotating, highlighting points of interest and annotating. This strongly suggests that the implemented gesture vocabulary does not need to allow full functionality of the PACS system and that many tools useful preoperatively for planning and diagnosis are unnecessary intraoperatively. As software generally faces a tradeoff between power and ease of use, these results indicate that future iterations should be geared toward making a stable system that focuses on basic access to and manipulation of images. Although using existing installed PACS software has advantages in terms of intraoperability and

familiarity, optimizing this software for use with gestural interfaces could prove beneficial.

CONCLUSION

With the success of this proof of concept and pilot series, further testing should be undertaken to determine how the intraoperative needs of different surgical specialties and subspecialties differ. This may, in turn, influence the choice of supported PACS features and the optimal gesture library. A larger study may also help quantify any time savings realized by the system and qualify its effects on surgical certainty and intraoperative education.

Competing interests: M. Strickland, J. Tremaine and G. Brigley formed the company GestSure Technologies based on the technology first described in this article. None declared for C. Law.

Contributors: All authors contributed to study design and approved the article's publication. M. Strickland, G. Brigley and C. Law acquired the data, which M. Strickland and C. Law analyzed. M. Strickland wrote the article, which J. Tremaine, G. Brigley and C. Law reviewed.

References

1. Luketich JD, Fernando HC, Buenaventura PO, et al. Results of a randomized trial of HERMES-assisted versus non-HERMES-assisted laparoscopic antireflux surgery. *Surg Endosc* 2002;16:1264-6.
2. El-Shallaly GEH, Mohammed B, Muhtaseb MS et al. Voice recognition interfaces (VRI) optimize the utilization of theatre staff and time during laparoscopic cholecystectomy. *Minim Invasive Ther Allied Technol* 2005;14:369-71.
3. Roe DB, Wang Y. A voice-controlled network for universal control of devices in the OR. *Minim Invasive Ther Allied Technol* 2000;9:185-91.
4. Alapetite A. Impact of noise and other factors on speech recognition in anaesthesia. *Int J Med Inform* 2008;77:68-77.
5. Punt MM, Stefels CN, Grimbergen CA, et al. Evaluation of voice control, touch panel control and assistant control during steering of an endoscope. *Minim Invasive Ther Allied Technol* 2005;14:181-7.
6. Graetzel C, Fong T, Grange S, et al. A non-contact mouse for surgeon-computer interaction. *Technol Health Care*. 2004;12:245-57.
7. Kipshagen T, Graw M, Tronnier V, et al. Touch- and marker-free interaction with medical software. In: *World Congress on Medical Physics and Biomedical Engineering*; 2009 Sept. 7-12; Munich, Germany. *IFMBE Proceedings* 2009;25:75-78.
8. Wachs J, Stern H, Edan Y, et al. Gestix: a doctor-computer sterile gesture interface for dynamic environments. In: Saad A, Avineri E, Dahal K, et al., editors. *Soft computing in industrial applications*. Berlin (Germany): Springer Berlin Heidelberg; 2007. p.30-9.
9. Wachs J, Stern H, Edan Y, et al. A gesture-based tool for sterile browsing of radiology images. *J Am Med Inform Assoc* 2008;15:321-3.

10. Wachs J, Stern H, Edan Y, et al. Real-time hand gesture interface for browsing medical images. *Int J Intel Comp Med Sci Image Proc* 2007;1: 175-85.
11. Cleary K, Peters T. Image-guided interventions: technology review and clinical applications. *Annu Rev Biomed Eng* 2010;12:119-42.
12. Hassfeld S, Mühling J. Computer assisted oral and maxillofacial surgery — a review and an assessment of technology. *Int J Oral Maxillofac Surg* 2001;30:2-13.
13. Nishikawa A, Hosoi T, Koara K, et al. FAcce MOUSe: a novel human-machine interface for controlling the position of a laparoscope. *IEEE Transactions on Robotics and Automation*. 2003;19:825-41.
14. Tani BS, Maia RS, Wangenheim AV. A gesture interface for radiological workstations. In: *Twentieth IEEE International Symposium on Computer-Based Medical Systems, 2007*; 2007 June 20-2; Maribor, Slovenia. New York: IEEE; 2007. p. 27-32.
15. Gallo L. A glove-based interface for 3D medical image visualization. In: Tsihrintzis GA, Damiani E, Virvou M, editors. *Intelligent Interactive Multimedia Systems and Services*. Smart Innovation, Systems and Technologies, Vol. 6. Berlin (Germany): Springer Berlin Heidelberg; 2010. 2010;6:221-30.
16. Pogue D. Kinect pushes users into a sweaty new dimension. *The New York Times* 2010 Nov. 4; B1.
17. Kinect hacked days after release. *BBC News website* 2010 Nov. 12. Available: www.bbc.co.uk/news/technology-11742236 (accessed 2011 Oct. 12).
18. Stern HI, Wachs J, Edan Y. Optimal consensus intuitive hand gesture vocabulary design. In: *Proceedings of the 2008 IEEE International Conference on Semantic Computing*; 2008 Aug. 4-7; Santa Clara, Calif. Washington: IEEE Computer Society; 2008. p. 96-103.
19. AORN Recommended Practices Committee. Recommended practices for maintaining a sterile field. *AORN J* 2006;83:402-4, 407-10, 413-16.
20. Wachs JP. Gaze, posture and gesture recognition to minimize focus shifts for intelligent operating rooms in a collaborative support system. *Int J Comput Commun Control* 2010;5:106-24.
21. Johnson R, O'Hara K, Sellen A, et al. Exploring the potential for touchless interaction in image-guided interventional radiology. In: *ACM Conference on Computer-Human Interaction (CHI)*; 2011 May 19-23; Vancouver, Canada. New York: Association for Computer Machinery; 2011. p. 3323-32.
22. Mathes AM, Kreuer S, Schneider SO, et al. The performance of six pulse oximeters in the environment of neuronavigation. *Anesth Analg* 2008;107:541-4.
23. van Oostrom JH, Mahla ME, Gravenstein D. The Stealth Station Image Guidance System may interfere with pulse oximetry. *Can J Anesth* 2005;52:379-82.
24. Hardware Info. *Open Kinect*. Available: http://openkinect.org/wiki/Hardware_info (accessed 2011 Nov. 2).