Clinical Evaluation of a New Touchless Interface That Uses Hand-Gestures and Voice Operation Named "Opect3D" in Cerebral Blood Vessel Cases

著者名	CHIBA Shinji, YOSHIMITSU Kitaro, ISHIKAWA Tatsuya, MASAMUNE Ken, HORISE Yuki, MURAGAKI Yoshihiro
journal or publication title	Tokyo Women's Medical University Journal
volume	4
page range	44-51
year	2020-12-25
URL	http://hdl.handle.net/10470/00032745

TWMUJ 4: 44-51, 2020

Original

Clinical Evaluation of a New Touchless Interface That Uses Hand-Gestures and Voice Operation Named “Opect3D” in Cerebral Blood Vessel Cases

Shinji Chiba,[1,2] **Kitaro Yoshimitsu,**[1] **Tatsuya Ishikawa,**[3]
Ken Masamune,[1] **Yuki Horise,**[1] and **Yoshihiro Muragaki**[1,3]

[1]Faculty of Advanced Techno-Surgery (FATS), Institute of Advanced Biomedical Engineering & Science, Graduate School of Medicine, Tokyo Women’s Medical University, Tokyo, Japan
[2]National Technology Office, Microsoft Japan Co., Ltd., Tokyo, Japan
[3]Department of Neurosurgery, Neurological Institute, Tokyo Women’s Medical University, Tokyo, Japan

(Accepted April 24, 2020)
(Advance Publication by J-STAGE June 16, 2020)

Background: In neurovascular therapy, normally surgeons directly instruct a radiologist to operate a computer on their behalf to maintain sterility in the operative field when reconfirming medical images. This process, however, takes time and causes psychological burden.

Methods: A touchless interface operated by hand-gestures and voice named “Opect3D” was developed to resolve the above issues. Initially, the Opect3D and conventional verbal communication were compared preclinically. The operation time for deciding the rotation angle and the angular difference against the targeted position were measured and calculated, respectively. Afterwards, Opect3D was utilized in unruptured cerebral aneurysms and the usefulness and psychological burden were surveyed.

Results: The mean time using the verbal communication was 24.1 ± 7.9 sec, while using Opect3D was 12.8 ± 5.3 sec per single interaction ($p < 0.01$). The angular difference using the verbal communication was 9.0 ± 8.4 degrees and 9.4 ± 6.1 degrees using Opect3D ($p = 0.87$). The survey of usefulness and psychological burden of using Opect3D yielded 4.03 ± 0.73 ($p < 0.02$) and 4.67 ± 0.54 ($p < 0.01$), respectively, as a mean value. Scores ranged from 1(worst) to 5(best), with 3 corresponding to the verbal communication. Significance was evaluated from the score of 3.

Conclusions: Usage of Opect3D was effective in reducing operation time and stress, thus enhancing efficiency and ease of operation.

Key Words: neurovascular therapy, three-dimensional, touchless, hand-gestures, voice control

Introduction

Neurovascular therapy is a method of choice for the treatment of cerebral aneurysms.[1] Coil embolization using a specialized catheter is one of several procedures used in neurovascular therapy.[2] The number of reports

Corresponding Author: Yoshihiro Muragaki, Faculty of Advanced Techno-Surgery (FATS), Institute of Advanced Biomedical Engineering & Science, Graduate School of Medicine, Tokyo Women’s Medical University, 8-1 Kawada-cho Shinjuku-ku Tokyo, 162-8666, Japan. ymuragaki@twmu.ac.jp
doi: 10.24488/twmuj.2020001

describing the effectiveness of neurovascular therapy have increased; accordingly, the number of interventions performed has also steadily grown.[3] In neurovascular therapy, accurate maneuvering of a therapeutic catheter is the key to successful treatment and enables the operator to smoothly reach lesions and maintain optimal contact with narrow vessel walls throughout the coil or stent.

In efforts to increase catheter maneuverability, surgeons can use three-dimensional model viewers, which enable them to accurately and precisely observe the target point from any angle. By observing the target relative to its surrounding tissues, the surgeon can decide on an appropriate rotation angle before treatment. When guiding a catheter to a cerebral aneurysm, the angle on the three-dimensional model viewer can be manipulated to aid in confirming the shape and angle of complex cerebral blood vessels. However, because surgeons wear sterilized gloves within the operative field, they cannot directly manipulate the models using a computer mouse or keyboard. Generally, manipulation of models is achieved through verbal communication between surgeons inside the operating theater and an externally located radiologist. Though they observe the same model, manipulative misunderstanding can occur because the conventional method using verbal communication and manual computer controls does not provide enough context for the radiologist. In such cases, surgeons may have to frequently leave the operating room to go to the radiologist's console room and indicate the desired manipulation. However, leaving the operative field during a procedure causes stress to the surgeon due to interruption of "surgical rhythm".

To resolve these issues, Iannessi et al. and Hotker et al. proposed a touchless intraoperative display system that facilitates several endovascular therapies, and attempted its use in a clinical setting.[4-6] Other trials using the Kinect sensor (Microsoft Corporation, Redmond, WA, USA) have been performed by Wipfli, Hochman, and a group led by Tan.[7-9] The Leap Motion™ system (Leap Motion, Inc., San Francisco, CA, USA), which uses infrared depth sensors, has been used in clinical trials and reported to be effective in clinical use.[10-13] A similar finding was reported in a gesture control study by Strickland et al.[14] However, in most cases of touchless interface systems, operating surgeons are forced to manipulate the model using both hands, which is not compatible with clinical use because surgeons are often required to handle other devices or catheters. As Yoshimitsu et al. reported in their description of the "Opect" system, surgeons need a system of manipulation that only requires the use of one hand.[15] The Opect was designed to display a two-dimensional image from a single point of view, and thus the operator need to repeat the same gesture to confirm the depth direction of the image. Previous models using the state-of-the-art communication tools described above inspired us to design a new communication interface that enables three-dimensional cerebrovascular model controllability with use of one hand and voice recognition. This article describes a touchless three-dimensional cerebrovascular model interface named "Opect3D" which enables an operating surgeon to manipulate three-dimensional models by using intuitive finger-gestures such as "pinching and rotating". Here, a prototype communication interface, which uses a newly designed touchless control algorithm, and experiences using the interface in clinical cases, is described.

Material and Methods

1. System configuration

The hand-gesture operation system uses the following hardware and software, the Opect3D, physical display for three-dimensional cerebrovascular model, a small personal computer (NUC, Intel Corporation, Santa Clara, CA, USA), a camera (RealSense SR 300, Intel Corporation, Santa Clara, CA, USA), and a three-dimensional model viewer (ParaView, Kitware, Clifton Park, NY, USA). RealSense has an infrared image processing-based camera and dual array microphones that can track human finger movements in three-dimensional spaces and can support voice control. The Opect3D takes over mouse functions. Therefore, the mouse does not need to be connected to the computer. Note this requires the installation of a free device driver. The operator can easily use it simply by bringing these sets of systems into the operating room, connecting them to the monitor, and turning on the PC. A processing flow of the Opect3D is shown in **Figure 1**.

2. Specification of Opect3D using RealSense

The farthest tracking distance of the Kinect sensor used in the Opect is 4.0 m, compared to 1.5 m for RealSense. While this ability is comparatively low in RealSense, for our purposes RealSense was found to be superior in finger tracking because the Kinect uses full body tracking and is consequentially less precise in recognizing fine finger movements. The tracking distance of the RealSense camera covers the range in which the operator can operate the Opect3D without leaving the operative field. Additionally, because the field of view for RealSense is approximately 25.7% wider than that of Kinect, a wider range can be sensed at the same distance. Furthermore, the RealSense camera operates at up to 60 frames/sec, compared to a max of 30 frames/sec for Kinect, therefore the hand movements of the operator are tracked more smoothly. Although Leap Motion, used in previous studies,[10-13] also boasts high frames/sec, its farthest tracking distance is extremely short (only 0.6 m). As such, this requires the surgeon to either move closer to the device, or deploy a non-sterile device in close proximity to the operating field; therefore, the Leap Motion is generally not suitable for use within the sterilized operative field.

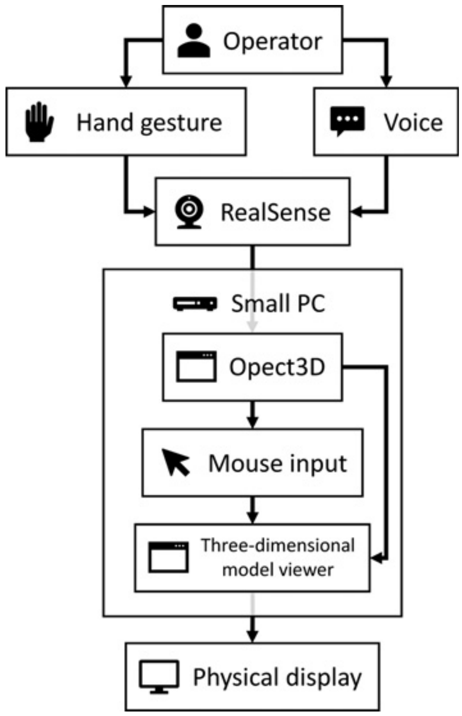

Figure 1 Flow diagram of the Opect3D interface. Hand-gestures and voice commands are detected by the RealSense camera and then fed into the Opect3D program in a small computer. The Opect3D takes over mouse function to control the three-dimensional model viewer. Three-dimensional medical images are generated by the model viewer on a physical display, and the viewing angle of those can be adjusted by hand-gestures and voice commands through the Opect3D.

3. Hand-gestures manipulation of Opect3D

During surgery, only the surgeon with operational authority is tracked by the Opect3D. Operational authority is obtained by the surgeon pointing to RealSense with their right hand and gesturing in a clockwise motion twice. If another member of the surgical team, for example a surgical assistant, wishes to assume operational authority, they can do so by performing the same gesture. **Table 1** lists the gesture commands available in the Opect3D. As a supplemental function, a voice control can be used to prevent malfunction during operation and to enlarge the model image. Moreover, since ParaView covers an image re-opening function, the surgeons can switch aneurysm models using their hands without leaving the operating field or compromising sterility.

Table 1 The surgeon can control Opect3D through touchless operation. There are five types of preset hand gestures and three types of voice commands. Voice commands are customizable to user preference.

Input type	Operation	Behavior
Gesture	Point to RealSense	Engage with the Opect3D.
Gesture	Turn one hand clockwise twice	Acquire operation authority.
Gesture	Pinch with thumb and index finger	Click the left mouse button.
Gesture	Release each finger	Release the mouse button.
Gesture	Move one hand while pinching	Rotate or translate the working angle. Draw annotations while screen zoom.
Voice	"Opect, stop"	Pause gesture input.
Voice	"Opect, restart"	Resume gesture input.
Voice	"Zoom"	Toggle screen zoom.

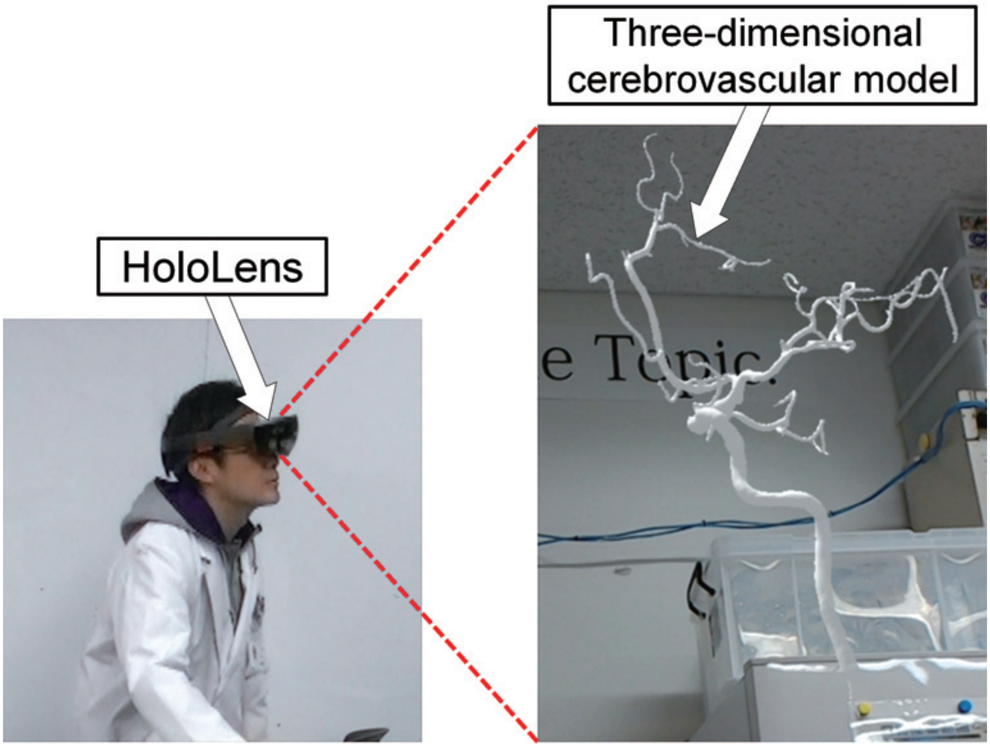

Figure 2 Setting of a targeted viewing angle using a head-mounted display; HoloLens in preclinical experiments. With the HoloLens, surgeons can observe a three-dimensional cerebrovascular model from various angles by changing their own perspective, and they can easily find their ideal perspective.

4. Preclinical evaluation of the verbal communication and Opect3D

In the experiment by Yoshimitsu et al.,[15]-the developers of Opect-there was no evaluation of its effectiveness compared with verbal communication. Accordingly, the authors of the present study believed that quantitative assessment of the results should be performed. However, to perform both the verbal communication and Opect3D under exactly the same conditions during clinical procedures is difficult. Furthermore, the final image decision for the rotation angle varies depending on the subjective opinion of the surgeon and the progress of the operation. Therefore, to set experimental conditions to be as uniform as possible, and to conduct proper evaluation of the system itself, two sets of preclinical experiments were conducted by 10 surgeons using a head-mounted display HoloLens (Microsoft Corporation, Redmond, WA, USA) (**Figure 2**) and performed as described the following steps. The HoloLens enables the surgeon to observe three-dimensional cerebrovascular model, that is constructed in a virtual space, as if it were floating in the air. The surgeon can freely observe the medical image in three-dimensions, and also can intuitively determine the target rotation angle.

Step 1. By equipping the HoloLens, the surgeon decides the rotation angle that they wish to view during surgery from the three-dimensional cerebrovascular model appearing in front of them and then remembers the model from an ideal perspective. Therefore, the rotation angle determined here is considered to be the correct target angle in the next step. The numerical data of the target angle was not indicated to the surgeons.

Step 2. Based on the image remembered by the surgeon in step 1, the rotation angle of the three-dimensional cerebrovascular model shown on the display was altered using both the verbal communication and Opect3D.

In the preclinical experiment for the verbal communication, the surgeon communicated to the radiologist verbally. The surgeon essentially operates the computer indirectly by issuing instructions to manipulate the displayed three-dimensional cerebrovascular model to approach the desired target angle. In the preclinical experiment for the Opect3D, the surgeon used the Opect3D to

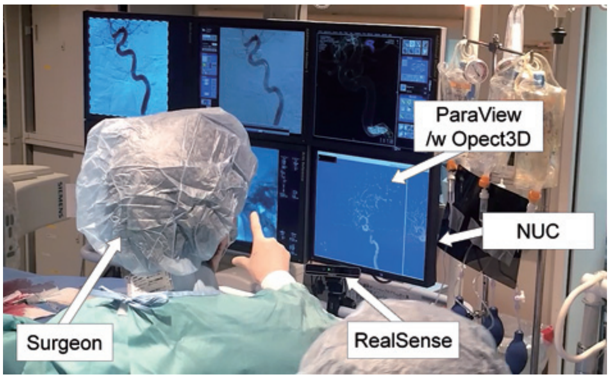

Figure 3 Clinical evaluation layout. RealSense is fixed to the display hanger and ParaView with the Opect3D is running. Of the numerous displays, the lower right display shows the outputted three-dimensional cerebrovascular model manipulated through touchless hand-gestures by the operator. The small computer (NUC) and keyboard are attached to the back of the display.

directly approach the target rotation angle through touchless operation. In the experiments, the operation time and angular difference between the targeted images and those determined by using the HoloLens were measured and calculated, respectively. The operation time is defined from started to operate an image to determine the viewing angle, and it was measured by the computer from event notifications of the mouse. In the first trial the verbal communication is followed by the Opect3D. In second trial the order is reversed and a separate medical image is used. Each surgeon performed the trials as described above.

5. Clinical evaluation of the Opect3D utility

The Opect3D was installed in the hybrid angiography and surgery rooms in the neurosurgical department of the Tokyo Women's Medical University. The Opect3D was used in 8 representative clinical cases of unruptured cerebral aneurysms that were treated with coil embolization to control the rotation angle and confirm the lesion site **(Figure 3)**. Because there was no direct intervention or materials used on the patients in this study, the ethics committee waived all requirements for review and informed consent. In the experiment, the RealSense camera was attached to the display hanger and was placed approximately 1.1 m in front of the surgeon to ensure no unintentional contact between the surgeon, instruments, and the non-sterile RealSense device. ParaView was used to generate high-quality three-dimensional model in real time by volume rendering (layered expression of slice images) a two-dimensional medical image in the Digital Imaging and Communications in Medicine (DICOM) format. A satisfaction survey querying three domains for the usefulness (rated 1 to 5, with 5 being the best) was administered to a total of 20 surgeons who used the Opect3D during surgery. The rated scores were compared with those in the verbal communication regarded as 3:

・Utility (Is the system able to support surgical decision making?)

・Functionality (Does the system function adequately?)

・Versatility (Does it appear that the system will be able to support future surgeries through cooperation with other devices and applications?)

Additionally, a questionnaire surveying three domains for psychological burden (rated 1 to 5) was also administered, as follows:

・Stress reduction (Does using the Opect3D reduce psychological stress compared to using the verbal communication?)

・Compromise reduction (Does the Opect3D reduce the need to use unsatisfactory images?)

・Significance (Whether the Opect3D is clinically significant?)

Results

1. Preclinical evaluation of the verbal communication and Opect3D

In terms of operation time per single interaction, the verbal communication took, on average, 24.1 ± 7.9 sec, while the Opect3D took 12.8 ± 5.3 sec ($p < 0.01$), with median values of 23.8 sec and 11.9 sec, respectively **(Figure 4A)**. These results suggest that the Opect3D requires approximately half the operation time compared to the verbal communication, which is a statistically significant difference in determining the appropriate rotation angle. The angle determined by the verbal communication was 9.0 ± 8.4 degrees off of the target angle, while the angle determined by the Opect3D was 9.4 ± 6.1 degrees off of the target angle; this difference, was not found to be statistically significant ($p = 0.87$) **(Fig-**

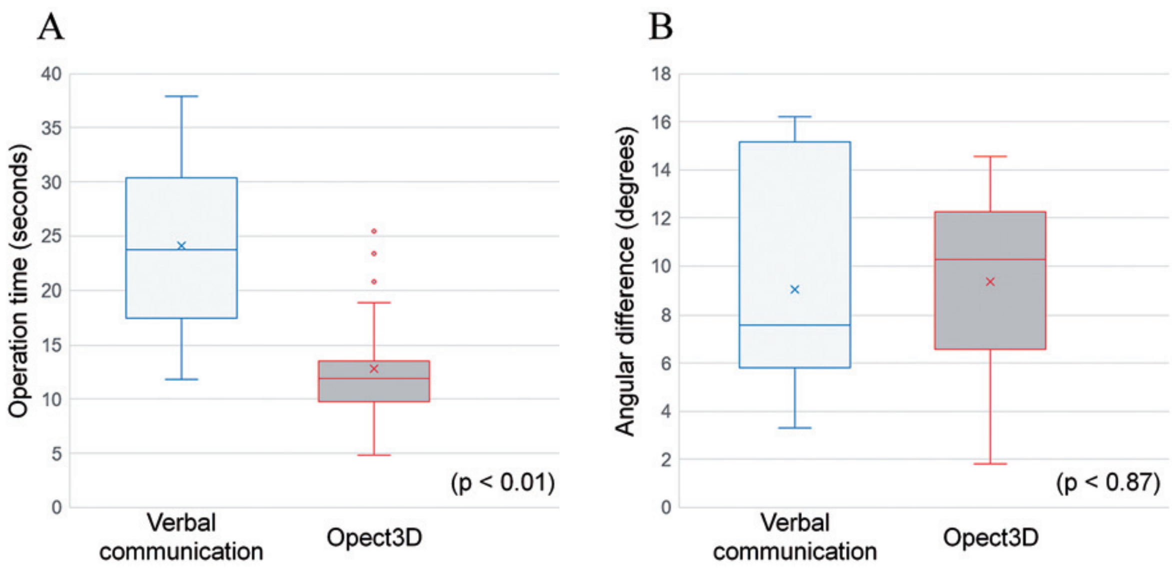

Figure 4 (**A**) The time required to change the viewing angle per interaction. The operation time using the Opect3D was significantly shorter compared to using the verbal communication. In one case of unruptured cerebral aneurysms, the surgeons needed to reconfirm the three-dimension medical image numerous times, thus the more often surgeons need to change the viewing angle, the more effective Opect3D is compared to the verbal communication.
(**B**) The angular difference between the surgeon's ideal and determined perspectives using both methods. There was no significant difference between the verbal communication and Opect3D.

ure 4B). This result supports that the Opect3D is effective.

2. Clinical evaluation of the Opect3D utility

Survey scores ranged from 1 being the worst to 5 being the best with 3 corresponding to the verbal communication, and the significance was evaluated from 3. Survey results in the usefulness of using the Opect3D in the clinical setting yielded mean scores of 3.95 ± 0.59 for utility, 3.55 ± 0.67 for functionality, and 4.60 ± 0.49 for versatility, with an overall average evaluation score of 4.03 ± 0.73 ($p < 0.02$) **(Table 2A)**. Survey results in the psychological burden yielded mean scores of 4.70 ± 0.46 for the stress reduction, 4.60 ± 0.66 for decrease in compromise and 4.70 ± 0.46 for clinical significance, with an overall average evaluation score of 4.67 ± 0.54 ($p < 0.01$) **(Table 2B)**.

Table 2 Survey results of using the Opect3D showing usefulness and psychological burden. Rating go from 1 to 5, with 5 being the best. A score of 3 indicates the verbal communication.

A.	Usefulness	
	Utility	3.95 ± 0.59
	Functionality	3.55 ± 0.67
	Versatility	4.60 ± 0.49
	Overall average score	4.03 ± 0.73
B.	Psychological burden	
	Stress reduction	4.70 ± 0.46
	Compromise reduction	4.60 ± 0.66
	Significance	4.70 ± 0.46
	Overall average score	4.67 ± 0.54

Discussion

1. Clinical usefulness of the Opect3D

Results of this experiment suggest that the Opect3D was more effective than the verbal communication in practical application. In the verbal communication, time loss was found to occur most frequently during communication between the surgeons and the radiologist. As such, the verbal communication required more time. Synchronization between the surgeon's intention and the Opect3D, contributed to reduced operation time compared to the verbal communication, which only vaguely conveyed the desired angular change. Compared to other studies using Kinect and Leap Motion, the Opect3D using RealSense can detect not only hand movements but also finger movements and voices. This superiority contributed to the benefit of time cost that the operation time by using the Opect3D was 1.9 times faster than that of the verbal communication.

From subjective evaluations from the surgeons and the radiologist, both sides agree that it was difficult to find the ideal perspective of viewing the three-dimensional cerebrovascular model using verbal communication. Both the radiologist and the surgeons felt that the psychological burden was reduced when using the Opect3D. The Opect3D was high rated in the survey because it reduced conflict and stress produced by miscommunication between surgeons and the radiologist.

2. Deployment of the Opect3D for smooth interventions

The Opect3D scored highly in the survey evaluating its versatility. Surgeons commented that the five types of preset finger gestures were easy to remember, they were able to manipulate the computer while maintaining the sterility without difficulty, and the three voice commands, including the zoom function, were also useful in clinical interventions. Surgeons did not complain of time lag, since they operated the Opect3D in real time. Although, incorrect recognition was observed during the experiment, surgical procedures were not affected by Opect3D operation. The Opect3D could be useful for intraoperative touchless image manipulation, in other operations such as surgery for intracranial arteriovenous malformations or installation of a ventriculoperitoneal shunt. The Opect3D is able to correlate movement of the operator's hand to the mouse pointer of the computer. Therefore, the Opect3D is available to be used for other three-dimensional model viewers. It could also be used for three-dimensional model viewers in fields outside of those used in this study. In the future, integration of artificial intelligence (AI) technologies to Opect3D is planned. This may enable smoother operation which will allow AI to indicate an optimal rotation angle of a medical image, calculated from previous image manipulation data performed by expert surgeons.

3. Limitation of the Opect3D in combination with the medical devices/software

Unfortunately, the Opect3D cannot currently be installed into legally approved medical devices such as picture archiving and communication systems (PACS), electronic medical recording systems (EMR), and planning software. Moreover, the RealSense cannot be used with various medical devices due to USB connection restrictions. Therefore, in this experiment, medical images captured during surgery could not be displayed, thus medical images exported from previous examinations were used. During clinical cases, the surgeon sometimes wanted to confirm the position of the catheter in the blood vessel using the Opect3D. To get around this restriction, the surgeon rotated the previously taken outpatient image using Opect3D then the radiologist came into the operating room to confirm the desired angle. The radiologist then produced the intraoperative image with the same view angle using conventional software. During this study the functionality of ParaView was also limited for the purpose of this study. Some surgeons commented that they would like to use Opect3D with numerical rotation angles and to use it for measuring the diameter and the volume of the aneurysms, these comments highlight the future potential of using Opect3D with a highly functional three-dimensional model viewer. The source code for Opect3D is planned to be license-free to be utilized so that many developers can integrate the Opect3D into various types of medical image viewers.

Conclusion

A touchless interface operated by hand-gestures and voice named Opect3D was developed. The Opect3D is compatible with any three-dimensional model viewers, for reconfirming medical images during surgery. The Opect3D resolved issues caused by miscommunication between surgeons and the radiologist during unruptured cerebral aneurysm operations. The results of the preclinical study suggest that the usability of the Opect3D was significantly superior to the verbal communication, due to reduce time during communication between surgeons and the radiologist. High ratings in satisfaction and subjective evaluations from surgeons and the radiologist showed that the Opect3D could reduce conflicts and stress between surgeons and the radiologist during operations. Therefore, the Opect3D was found to be an effective interface that can enhance the efficiency and quality of surgeries.

Acknowledgements

We would like to thank the following individuals for their guidance and assistance in the creation of this article and for their invaluable cooperation in the research, development and experimentation; Dr. Hiroshi Iseki, Dr. Nobuhisa Hagiwara, Dr. Hiroyuki Sakurai, Dr. Takakazu Kawamata, Dr. Tatsuya Shimizu, Dr. Takashi Maruyama, Dr. Bikei Ryu, Dr. Masao Usukura, and Dr. Yoshiyuki Konishi of Tokyo Women's Medical University, Dr. Tomoyuki Yano and Dr. Hideaki Oda of the Cancer Institute Hospital of JFCR, Dr. John Okcuoglu of Koc University School of Medicine, Dr. Takeyoshi Dohi of Tokyo Denki University, Mr. Yasuyuki Higuchi of the Panasonic Corporation, and all the people concerned at Microsoft Japan. This research was partly supported by research grants from "Subsidies for Establishment of Research Centers (Human Resources Project for Producing Advanced Medical Innovation)", "New Paradigms - Establishing Centers for Fostering Medical Researchers of the Future" and "Fostering Human Resources for Producing Medical Innovation".

Conflicts of Interest: All authors have no conflict of interest to declare

References

1. Liebman KM, Severson MA 3rd: Techniques and devices in neuroendovascular procedures. Neurosurg Clin N Am 20: 315–340, 2009
2. Taki W: Memorial review celebrating the 50th year of publication of NMC--neuroendovascular therapy. Neurol Med Chir (Tokyo) 50: 809–823, 2010
3. Dumont TM, Eller JL, Mokin M et al: Advances in endovascular approaches to cerebral aneurysms. Neurosurgery 74: S17–S31, 2014
4. Iannessi A, Marcy PY, Clatz O et al: Touchless intraoperative display for interventional radiologist. Diagn Interv Imaging 95: 333–337, 2014
5. Iannessi A, Marcy PY, Clatz O et al: Touchless user interface for intraoperative image control: almost there. Radiographics 34: 1142–1144, 2014
6. Hötker AM, Pitton MB, Mildenberger P et al: Speech and motion control for interventional radiology: requirements and feasibility. Int J Comput Assist Radiol Surg 8: 997–1002, 2013
7. Wipfli R, Dubois-Ferrière V, Budry S et al: Gesture-Controlled Image Management for Operating Room: A Randomized Crossover Study to Compare Interaction Using Gestures, Mouse, and Third Person Relaying. PLoS One 11: e0153596, 2016
8. Hochman JB, Unger B, Kraut J et al: Gesture-controlled interactive three dimensional anatomy: a novel teaching tool in head and neck surgery. J Otolaryngol Head Neck Surg 43: 38, 2014
9. Tan JH, Chao C, Zawaideh M et al: Informatics in Radiology: developing a touchless user interface for intraoperative image control during interventional radiology procedures. Radiographics 33: E61–E70, 2013
10. Park BJ, Jang T, Choi JW et al: Gesture-Controlled Interface for Contactless Control of Various Computer Programs with a Hooking-Based Keyboard and Mouse-Mapping Technique in the Operating Room. Comput Math Methods Med 2016: 5170379, 2016
11. Bizzotto N, Costanzo A, Bizzotto L et al: Leap motion gesture control with OsiriX in the operating room to control imaging: first experiences during live surgery. Surg Innov 21: 655–656, 2014
12. Rosa GM, Elizondo ML: Use of a gesture user interface as a touchless image navigation system in dental surgery: Case series report. Imaging Sci Dent 44: 155–160, 2014
13. Mewes A, Saalfeld P, Riabikin O et al: A gesture-controlled projection display for CT-guided interventions. Int J Comput Assist Radiol Surg 11: 157–164, 2016
14. Strickland M, Tremaine J, Brigley G et al: Using a depth-sensing infrared camera system to access and manipulate medical imaging from within the sterile operating field. Can J Surg 56: E1–E6, 2013
15. Yoshimitsu K, Muragaki Y, Maruyama T et al: Development and initial clinical testing of "OPECT": an innovative device for fully intangible control of the intraoperative image-displaying monitor by the surgeon. Neurosurgery 10 (Suppl): 46–50, 2014