



# Techniques for developing and viewing stereoscopic three-dimensional teaching videos for transoral robotic surgery (TORS)

Ernest D. Gomez<sup>1</sup> · Erica R. Thaler<sup>1</sup> · Bert W. O'Malley Jr.<sup>1</sup> · Christopher H. Rassekh<sup>1</sup> · Gregory S. Weinstein<sup>1</sup> · Jason G. Newman<sup>1</sup> · Robert M. Brody<sup>1</sup>

Received: 19 March 2019 / Accepted: 23 March 2019 / Published online: 3 April 2019  
© Springer-Verlag London Ltd., part of Springer Nature 2019

## Abstract

Advances in three-dimensional (3D) video recording and playback have expanded the availability of stereoscopic videos for consumption with specialized televisions, computer monitors, and virtual reality (VR) headsets. The built-in stereoscopic vision of the da Vinci surgical system (Intuitive Surgical, Sunnyvale, CA) enables the recording of both camera views. However, medical-grade devices for direct stereoscopic video recording can be cost-prohibitive. We describe methods of producing and viewing 3D videos in transoral robotic surgery (TORS) without acquisition of dedicated 3D recording equipment. The video output from the left and right cameras of the da Vinci Si endoscope were recorded during a TORS radical tonsillectomy using non-stereoscopic media devices. Raw video was processed using commercially available video-editing software to create stereoscopic videos clips. 3D videos were successfully observed using a low-cost mobile phone VR headset. 3D stereoscopic surgical videos were successfully developed using widely available software, applications, computer peripherals, and mobile devices. The methods described confer significant cost savings when compared to the purchase of specialized recording equipment despite the requirements of time and intermediate computer skills. Future studies will investigate the utility of stereoscopic videos in medical student and resident education.

**Keywords** Three-dimensional · Virtual reality · Stereoscopic · Surgical video

## Introduction

Transoral robotic surgery (TORS) has become a valuable tool in otorhinolaryngology—head and neck surgery over the past decade. Initial prospective trials involving TORS began in 2005 and were followed by FDA clearance in 2009 [1–3]. The da Vinci surgical system (dVSS, Intuitive Surgical Inc, Sunnyvale, CA) provides head and neck surgeons with high definition three-dimensional visualization of the oropharynx and larynx, access to difficult to reach areas in the

upper aerodigestive tract, and multiple degrees of freedom of movement for surgical instrumentation.

Although the surgeon console provides depth perception through stereopsis, standard recording equipment and the vision cart only provide the bedside assistant and other observers with a two-dimensional (2D) view. Stereopsis, or depth perception, is provided to the surgeon console by displaying two slightly different images to each eye which are superimposed by the brain to produce a three-dimensional (3D) image [4].

3D video has been acquired for both open and endoscopic procedures using a variety of available devices. Prior investigators have captured open surgeries in 3D with side-by-side GoPro cameras (GoPro Inc, San Mateo, CA) [5] and smartphones [6]. Since all models of the dVSS include a 3D stereoscopic camera, the remaining barriers to 3D video production for robotic surgery are simultaneous left–right eye recording, image synchronization, and image alignment—functions normally automated by specialized 3D recording devices. In this article, we demonstrate that 3D videos can

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s11701-019-00952-z>) contains supplementary material, which is available to authorized users.

---

✉ Robert M. Brody  
Robert.brody@uphs.upenn.edu

<sup>1</sup> Department of Otorhinolaryngology, Head and Neck Surgery, Hospital of the University of Pennsylvania, 5th Floor Ravdin, 3400 Spruce Street, Philadelphia, PA 19104, USA

be produced and viewed without specialized 3D video equipment at a low cost.

## Methods

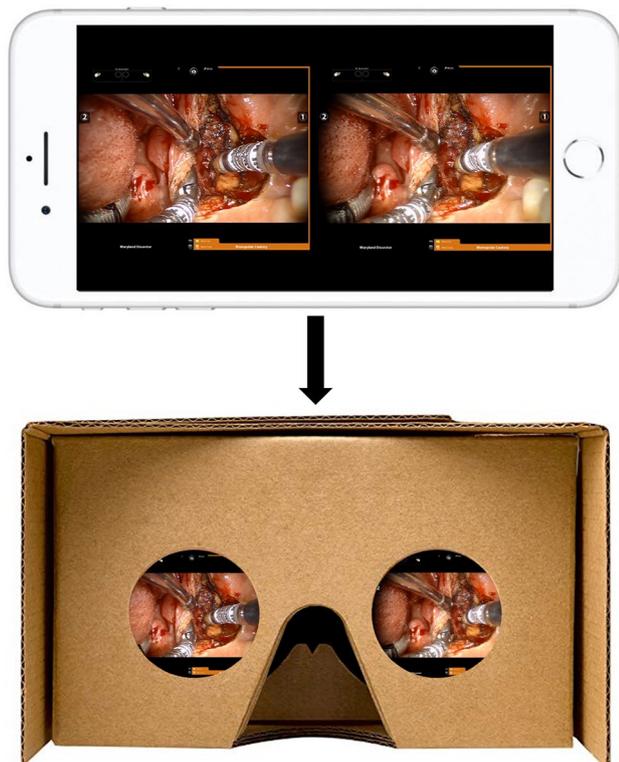
### Stereoscopic video recording

The vision cart of each dVSS model has video output ports for both the left and right camera feeds. Both video outputs must be recorded to produce a stereoscopic video. A fully 3D viewing capable video has two video streams corresponding to the left and right eye views. The two videos must be synchronized and positioned symmetrically about the center of a viewing screen. A still image of a 3D TORS video viewed on a smartphone is presented in Fig. 1.

Specialized 3D video recording units such as the Sony HVO-4000MT (Sony, Minato, Tokyo, Japan), the Karl Storz AIDA (Storz, Tuttlingen, Germany), and the Olympus IMH-20 (Olympus, Shinjuku, Tokyo, Japan) are capable of recording synchronized, side-by-side left and right eye videos that immediately ready for 3D viewing. However, these proprietary medical-grade equipments for 3D video recording can be costly, with video devices priced between

\$5000 and \$20,000. Other devices such as the Stryker SDC3 HD Information Management System (Stryker, Kalamazoo, Michigan, USA) are capable of simultaneous left and right view recording, but the two feeds are recorded in separate files and require video editing on a computer to synchronize the two images and position them side-by-side. Left and right video feeds can also be recorded on two separate devices, which can include any pair of the following: single-video medical-grade media recorders, digital video disk (DVD) recorders, or video game capture devices such as the Elgato HD60S (Elgato, Munich, Germany) and AVerMedia Live Gamer (AVerMedia, New Taipei City, Taiwan). It is not essential for the two recording devices to be identical, but final video quality and color consistency is maximized using the same device make and model for each video feed. DVD recordings must be converted to an editable digital format using a computer. Game capture devices require a dedicated computer in the operating room and a relatively complex setup compared to more integrated medical-grade systems. A summary of each method and its relative monetary and time costs is presented in Fig. 2.

In this study, two separate DVD recorders were used to record the left and right video feeds during a TORS radical tonsillectomy. An Institutional Review Board waiver was obtained by removing all identifiable information including metadata from the video. Video files from each DVD were extracted into a computer and converted to MPEG-4 format. Using Adobe Premiere (Adobe Inc., San Jose, CA) video-editing software, the left and right camera views were placed side-by-side. The two DVD recorders could not be started in perfect synchrony, so excess frames resulting from delay in recording start between the two DVD recorders were clipped using the software. The left and right video feeds were placed symmetrically about a middle vertical line in the video frame. This method was successfully replicated using Final Cut Pro (Apple, Cupertino, CA, USA) as well as Lightworks (EditShare, Watertown, MA), a freeware program.



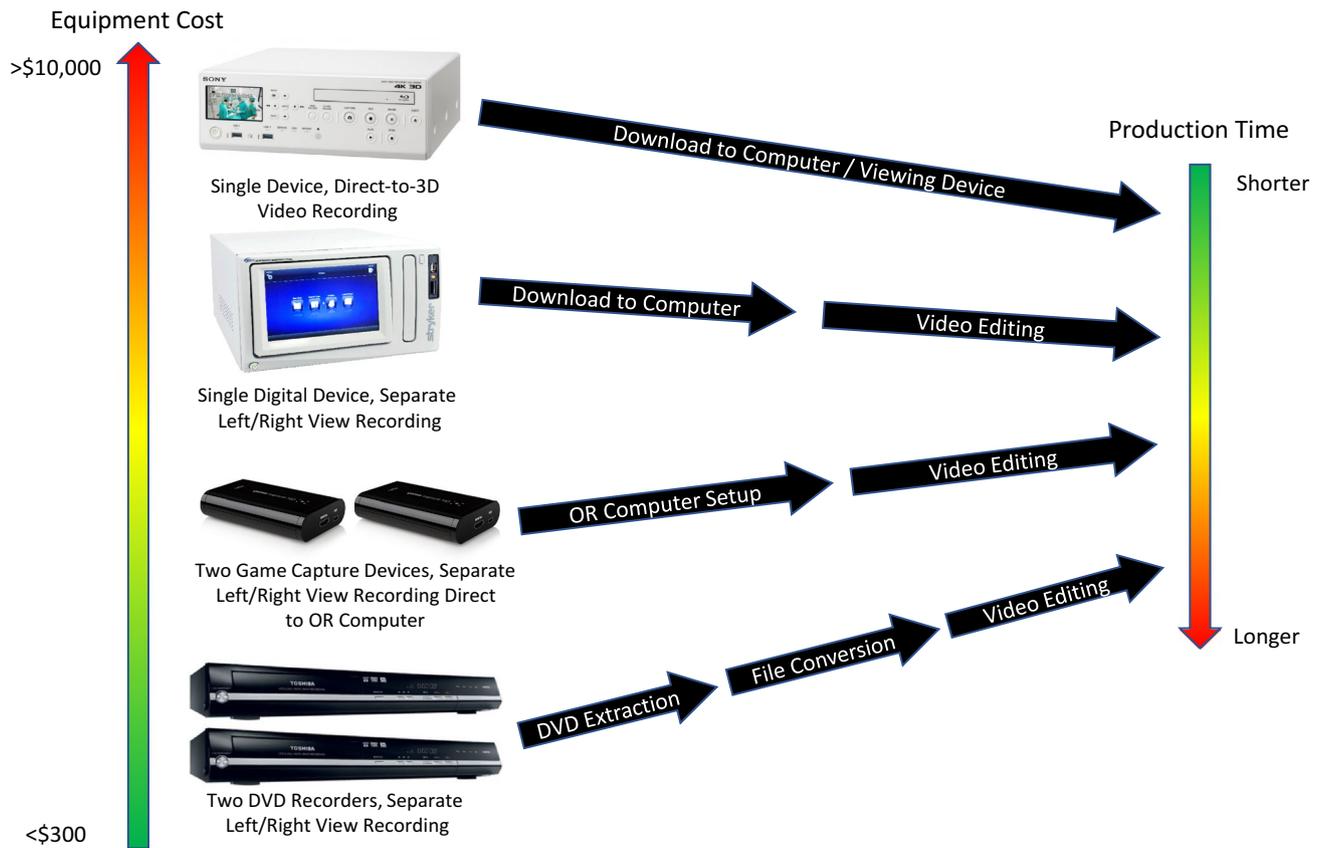
**Fig. 1** The left and right video feeds of a da vinci surgical system are juxtaposed to create a stereoscopic video which can be viewed on a smartphone (top). The video can then be viewed using low-cost virtual reality headsets such as the Google cardboard (bottom)

### 3D video viewing

3D videos can be viewed in anaglyph with red-cyan glasses, passive polarizing glasses, active shutter glasses, or stereoscopic VR headsets.

Videos intended for many users as in the case of a classroom can be uploaded to YouTube (Google, Menlo Park, CA), which has built-in capabilities for 3D viewing in both anaglyph and VR modes. The Google Cardboard (Menlo Park, CA) app and headset were used to view the TORS radical tonsillectomy 3D video on both iPhone (Apple, Cupertino, CA) and Android mobile phones (Fig. 1).

3D televisions can be used to view live videos in the operating room as well as recorded videos in the didactic setting. 3D televisions use passive polarizing glasses to provide



**Fig. 2** Flowchart of 3D video production methods. The figure demonstrates the time–cost trade-off of 3D surgical video production using various methods described in this work

the observer’s left and right eye with its respective camera view. Computer monitors equipped with active shutter glasses such as the NVIDIA 3D Vision System (NVIDIA, Santa Clara, CA) can also be used. Polarizing glasses require greater capital expenditures than the Google Cardboard VR headset but are viable options when cost is less prohibitive.

Anaglyph red-cyan glasses can be used to view 3D videos on mobile devices, 3D televisions, and computers, but the authors recommend against use of anaglyph glasses for learning surgical anatomy due to the color distortion from the tinted lenses.

## Results

Using 3D video recorded from the da Vinci Si and stereoscopic video-editing software, a 3D demonstration video for TORS radical tonsillectomy was developed. The lowest cost method for production (assuming no equipment was readily available) was achieved using two separate DVD recorders (cost of approximately \$200 each), the Lightworks video-editing software (free), and the Google Cardboard VR headset (\$12). A video clip of a TORS radical tonsillectomy

performed with the da Vinci Si Surgical System is available as an online supplement.

## Discussion and future directions

This study successfully demonstrates a low-cost technique for 3D stereoscopic robotic surgical video production for TORS. The methods presented in this work can be used for any robotic surgical procedure using the dVSS. Most centers with a surgical robot have access to at least one available device for video recording, so the marginal cost of producing and viewing 3D videos using the methods described can be up to three orders of magnitude lower than the cost of specialized 3D recording equipment.

3D videos afford a perspective not available by standard teaching videos. 3D vision has been associated with faster acquisition of laparoscopic and robotic technical skills among novices [7, 8]. Such findings support the value of depth perception in the performance of endoscopic and robotic tasks.

Furthermore, viewing 3D videos with a VR headset may simulate the experience of a console surgeon in a manner

that provides a more immersive learning environment than viewing a 2D screen. The value of 3D video has been reported for neurosurgical education [4, 9]. The integration of mobile devices into education has also been suggested as way to adapt medical education to millennial learners [10]. Additional studies during robotic surgical training are warranted to assess whether supplementary 3D videos prior to simulation or clinical practice improve knowledge retention and skill acquisition.

### Compliance with ethical standards

**Conflict of interest** Dr. Ernest D. Gomez has received research funding from Biomet and is a co-inventor on a patent for a robotic surgical device (US 9990856B2). Dr. Erica R. Thaler has received research funding from Inspire. Dr. Bert W. O'Malley, Jr., and Dr. Gregory S. Weinstein receive royalties from Olympus. Dr. Jason G. Newman is a medical consultant for Medtronic, Castle Biosciences, and Just Right Surgical. Dr. Christopher H. Rassekh and Dr. Robert M. Brody declare that they have no conflict of interest.

### References

- Weinstein GS, O'Malley BW Jr, Snyder W, Sherman E, Quon H (2007) Transoral robotic surgery: radical tonsillectomy. *Arch Otolaryngol Head Neck Surg* 133(12):1220–1226. <https://doi.org/10.1001/archotol.133.12.1220>
- O'Malley BW Jr, Weinstein GS, Snyder W, Hockstein NG (2006) Transoral robotic surgery (TORS) for base of tongue neoplasms. *Laryngoscope* 116(8):1465–1472. <https://doi.org/10.1097/01.mlg.0000227184.90514.1a>
- Weinstein GS, O'Malley BW Jr, Snyder W, Hockstein NG (2007) Transoral robotic surgery: supraglottic partial laryngectomy. *Ann Otol Rhinol Laryngol* 116(1):19–23
- Benet A, Tabani H, Griswold D, Zhang X, Kola O, Meybodi AT, Lawton MT (2016) Three-dimensional imaging in neurosurgical research and education. *World Neurosurg* 91:317–325. <https://doi.org/10.1016/j.wneu.2016.04.023>
- Birnbaum FA, Wang A, Brady CJ (2015) Stereoscopic surgical recording using GoPro cameras: a low-cost means for capturing external eye surgery. *JAMA Ophthalmol* 133(12):1483–1484. <https://doi.org/10.1001/jamaophthalmol.2015.3865>
- Gallagher K, Jain S, Okhravi N (2016) Making and viewing stereoscopic surgical videos with smartphones and virtual reality headset. *Eye (Lond)* 30(4):503–504. <https://doi.org/10.1038/eye.2015.282>
- Poudel S, Kurashima Y, Watanabe Y, Ebihara Y, Tamoto E, Murakami S, Nakamura T, Tsuchikawa T, Okamura K, Shichinohe T, Hirano S (2017) Impact of 3D in the training of basic laparoscopic skills and its transferability to 2D environment: a prospective randomized controlled trial. *Surg Endosc* 31(3):1111–1118. <https://doi.org/10.1007/s00464-016-5074-8>
- Wagner OJ, Hagen M, Kurmann A, Horgan S, Candinas D, Vorburger SA (2012) Three-dimensional vision enhances task performance independently of the surgical method. *Surg Endosc* 26(10):2961–2968. <https://doi.org/10.1007/s00464-012-2295-3>
- Sorenson J, Khan N, Couldwell W, Robertson J (2016) The rhoton collection. *World Neurosurg* 92:649–652. <https://doi.org/10.1016/j.wneu.2016.06.037>
- Toohy SL, Wray A, Wiechmann W, Lin M, Boysen-Osborn M (2016) Ten tips for engaging the millennial learner and moving an emergency medicine residency curriculum into the 21st century. *West J Emerg Med* 17(3):337–343. <https://doi.org/10.5811/westjem.2016.3.29863>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.